APPROXIMATION METHODS IN INDUCTIVE INFERENCE

By

WILLIAM R. MOSER II

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1996

UNIVERSITY OF FLORIDA LIBRARIES

For Maggie, who is happily impulsive, and decidedly unscientific in her approach to the world

ACKNOWLEDGEMENTS

First, my heartfelt thanks and appreciation go to Professor Douglas Cenzer for his kind help and patience in guiding me through the preparation of this dissertation. He taught me not simply the mathematical elements of recursion theory, but also instilled in me a sense of the style and philosophy with which the subject should be approached. Dr. Cenzer allowed me the time to find my own path into the area of inductive inference, and the opportunity to pursue it, and for this I am grateful. I would also like to thank Professors Glover, Mitchell, Smith and Wilson for taking the time to read this dissertation and to serve on my supervisory committee.

Finally, I am especially grateful to my parents for their encouragement, counsel, and support in my pursuit of this dream.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

APPROXIMATION METHODS IN
INDUCTIVE INFERENCE

By

William R. Moser II

August 1996

Chairman: Dr. Douglas A. Cenzer
Major Department: Mathematics

In many areas of scientific inquiry, the phenomena under investigation are viewed as functions on the real numbers. Since observational precision is limited, it makes sense to view these phenomena as bounded functions on the rationals. One may translate the basic notions of recursion theory into this framework by first interpreting a partial recursive function as a function on $\mathbb{Q}$. The standard notions of inductive inference carry over as well, with no change in the theory.

When considering the class of computable functions on $\mathbb{Q}$, there are a number of natural ways in which to define the *distance* between two functions. We utilize standard metrics to explore notions of *approximate* inference — our inference machines will attempt to guess values which converge to the correct answer in these metrics. We show that the new inference notions, NV_∞, EX_∞, and BC_∞, infer more classes of functions than their standard counterparts, NV, EX, and BC. Furthermore, we give precise inclusions between the new inference notions and those in the standard

inference hierarchy. We also explore weaker notions of approximate inference, leading to inference hierarchies analogous to the EX^n and BC^n hierarchies. Oracle inductive inference is also considered, and we give sufficient conditions under which approximate inference from a generic oracle G is equivalent to approximate inference with only finitely many queries to G.

Finally, we employ approximation techniques from topology and analysis to obtain a somewhat surpising result regarding triviality in oracle inductive inference classes. A result of Sacks states that if A is non-recursive, then the set $\{B \mid A \leq_T B\}$ has measure zero. Thus, from a measure theoretic perspective, a nontrivial (i.e. non-recursive) oracle is comparable to few oracles containing more information in the partial order of Turing degrees. We show that analogous results hold for the standard notions of inductive inference, as well as for the notions of approximate inference.

CHAPTER 1
INTRODUCTION

1.1 Basic Models of Inductive Inference

In the experimental sciences, it is typical to make generalizations from a finite set of data, amending such generalizations as needed to account for new data. Often, the data are collected at discrete time intervals and so can be viewed as ordered in a natural way. Also, the "experiment" need never end — we may continue to record a new datum during each time interval and make a new generalization at this time. It is this situation which we take as our starting point. As a further simplification we view such sequences of data as a functions on the natural numbers. In the next two sections, we give an informal account of the basic models of inductive inference, and some common variants. See Popper [19] for a philosophical perspective on notions of inductive inference. A more detailed mathematical treatment, giving the formal definitions and stating the major results of inductive inference, is found in chapter 2.

Gold [10] was among the first to formalize an example-based theory of learning. The two basic inductive inference paradigms we informally describe below are essentially due to him. The reader is referred to the excellent survey [17] by Odifreddi for a detailed introduction to the subject of inductive inference from a recursion theoretic perspective. In the simplest inference model, we wish to define a procedure which works as a predictor for a given class of phenomena. Given a sequence of data (a partial record of a phenomenon in the class under consideration), we wish to predict what will occur next. In this type of inference, called *next-value* inference, we use the sequence of data $\langle d_0, \ldots, d_{n-1} \rangle$ collected prior to time n to predict the value of the datum d_n which we will collect at time n. In view of Church's thesis, we will

assume that we must do so *algorithmically.* That is, our inference procedure must be a recursive function — we may not use coin flips, oracle consultations, appeals to higher beings, etc. in our attempt to predict the next datum.

We must now consider the accuracy requirements of our prediction algorithm. At time $t = 0$, we are hardly in a position to make a reasonable guess for d_0. If the underlying phenomenon is complex, we may not be in a position to make a reasonable guess for d_N at time $t = N$, even for large values of N. We will thus allow our prediction algorithm to make some mistakes, but we will require of the algorithm that it is *eventually* always accurate. In other words, for an algorithm to be considered predictive, it must make only finitely many errors in predicting the data. Henceforth, if a class $\mathcal{C}$ of phenomena is predicted by an algorithm M, we will say that M *NV*-identifies $\mathcal{C}$, or that $\mathcal{C} \in NV$.

Notice that this restriction on the error rate of our prediction algorithm essentially forces us to assume that the phenomena we are attempting to predict are also algorithmic, since our prediction machine eventually outputs the same values as those observed experimentally. Thus, we impose the further restriction on our class of phenomena that they may be viewed as *recursive* functions; hence, our "world" is the class of recursive functions.

Since we now view all phenomena as "black boxes" implementing algorithms of some sort, it is natural to use the data gathered from a given "black box" to attempt to guess its underlying algorithm. Instead of simply guessing the next output, we wish to *explain* what is going on inside the box. We still have only the data stream to work with; so, at time n we will again use the data gathered prior to this time to infer the underlying rule. As with the predictive machine, we allow the explanatory machine to make finitely many mistakes, as long as it eventually settles on an algorithm for each phenomenon in the class it explains. If a class $\mathcal{C}$ of recursive functions is explained by an algorithm M, we will say that M *EX*-identifies $\mathcal{C}$, or that $\mathcal{C} \in EX$.

Let us illustrate these concepts with a concrete example. The "phenomena" under consideration will be the class $\mathcal{P}$ of polynomials on the natural numbers. Consider the following machine M, which takes as input $\langle f(0), \ldots, f(n-1)\rangle$ at stage n (let $L(x) \equiv 0$):

Stage n:

Step 1: If $L(x) = f(x)$ for all $x < n$, output $L(n)$. Otherwise,

Step 2: Form the Lagrange interpolant,

$$L_n(x) = \sum_{i=0}^{n-1} f(i) \prod_{j \neq i} \frac{x-j}{i-j}.$$

Set $L = L_n$, and output $L(n)$.

End of Construction.

Now, if the phenomenon being observed is a polynomial p of degree n, we know that the $n+1$ distinct data points $\langle p(0), \ldots, p(n)\rangle$ uniquely determine this polynomial, so our algorithm will produce correct outputs after time $n+1$. Thus, our algorithm *NV*-identifies $\mathcal{P}$. It is similarly easy to construct a machine which *EX*-identifies $\mathcal{P}$: at time n, instead of outputting $L(n)$, output the interpolant L itself.

For another example, consider the class $\mathcal{PR}$ of primitive recursive functions. Since we may effectively enumerate the members of $\mathcal{PR}$, it is easy to construct a prediction machine for the class. Given inputs $\langle a_0, \ldots, a_{n-1}\rangle$ at time n, we simply search for the first function p in our enumeration which agrees with the given data, i.e. such that $p(0) = a_0, p(1) = a_1, \cdots p(n-1) = a_{n-1}$, and output $p(n)$. An explanatory machine would use a similar procedure, but output the function p at stage n. At this point, we might guess that the family *NV* of classes predicted by some algorithm is identical to the family *EX* of explanatory classes. The answer depends on what kinds of algorithms an explanatory machine is allowed to output as its guesses.

1.2 Variants of the Basic Models

In a scientific explanation of the phenomenon underlying observed data, it is usually the case that the explanation is *consistent* with the data. In view of this, we may require of our explanatory machine that the algorithm output at stage n agrees with the data given up to this time. In addition, the unsolvability of the Halting problem tells us that it is in general impossible to determine whether our output algorithms are total. Thus it is also natural to allow our explanatory machine to output algorithms which are consistent with the available data at time n, but which may not return an answer for times $m > n$ if we attempt to use the algorithm to predict such future values. If we denote the family of phenomena classes which can be explained in this manner by EX_{cons}, then EX_{cons} strictly contains NV. If, however, we restrict our explanatory machine to output only total algorithms, then the two families are equal.

Many other variations on the basic model have been explored. Blum and Blum [4] require that if an explanatory machine eventually stabilizes on an algorithm, then this algorithm accurately describes the phenomenon. This type of inference is known as *identification by reliable explanation*, and the family of identified classes as EX_{rel}. EX_{rel}-identification is more general than EX_{cons}-identification, but more restrictive than EX-identification. Feldman [8] requires only that the explanatory machine eventually output algorithms which are extensionally identical; the same algorithm need not appear twice, but the algorithms output should all produce identical outputs on identical inputs. This type of inference is known as *behaviorally correct identification*, and it subsumes EX-identification.

One may further relax the requirements for an explanatory machine by accepting explanations which come "close" to describing the observed phenomenon. Specifically, we may allow the algorithm upon which the explanatory machine finally

settles to differ finitely from the observed data. Furthermore, we may require a uniform upper bound on the number of errors allowed, or we may be a bit less restrictive, settling for identification with *arbitrary* finite errors. Case and Smith [6] show that the hierarchies of the induced inference families are proper and that behaviorally correct identification with arbitrary finite errors is powerful enough to identify all the recursive functions.

Many other variants appear in the literature. For example, Royer [22] defines notions of probablistic inference, and Case, Jain and Sharma [5] define notions of limiting inference which are quite different from the notions of approximate inference explored in the sequel.

1.3 Approximate Inference

Not all scientific inquiry requires such strong notions of inference. Often, the phenomenon under investigation is assumed to be a bounded continuous function on the real numbers. After receiving only a finite number of datum, the phenomenon is then represented by a function interpolating the known data. While the interpolant may *equal* the underlying function at only a few points, it is nevertheless viewed as a reasonable representation of this function. In the chapters that follow, we will formalize notions of inference arising from this point of view.

In chapter 3, we introduce the concept of inference of a recursive rational-valued function and define an important subclass of these functions, *RUC*, which cannot be inferred by the standard inference methods, but for which a natural approximation technique exists. This technique leads in chapter 4 to new "approximate" inference paradigms which subsume the standard paradigms and allow us to easily infer the class *RUC*. In chapter 5, we define hierarchies of approximate inference classes, the "epsilon" inference classes, which enable us to infer more classes than with the approximate inference classes, while retaining the finitistic flavor of the standard inference notions. In chapter 6, we use approximation techniques from topology and

analysis to obtain a surprising result regarding triviality of inference classes. Finally, we conclude in chapter 7 with some further directions for research in the area of approximate inference.

CHAPTER 2
PRELIMINARIES

2.1 Notions from Recursion Theory

Let $I_{\mathbb{Q}} = [0,1] \cap \mathbb{Q}$. We assume that $\{q_0, q_1, \dots, q_n, \dots\}$ is a fixed 1-1 effective enumeration of $I_{\mathbb{Q}}$ such that $q_0 = 0$, and $q_1 = 1$. $[I_{\mathbb{Q}}]^{<\omega}$ will denote the set of finite sequences of rationals. An arbitrary (partial) recursive function can be viewed as a function on $I_{\mathbb{Q}}$, by interpreting the natural number m as the m^{th} rational in the enumeration $\{q_n\}_{n\in\omega}$. More formally, we could define the *rational interpretation* $f_{\mathbb{Q}}$ of a (partial) recursive f by $f_{\mathbb{Q}}(q_m) = q_n$ iff $f(m) = n$. In this framework, we could carry out all basic recursion theory, proving the enumeration theorem, the S^m_n and fixed point theorems, and so forth. In the sequel, we will dispense with most of this formalism, preferring to use notation to indicate the types of functions under consideration.

Let $\langle a_0, \dots, a_{n-1}\rangle$ denote the usual coding of a finite sequence of natural numbers (or rationals) by a natural number. For a function f and natural number n, $f|_n$ denotes $\langle f(0), \dots, f(n)\rangle$ (or $\langle f(q_0), \dots, f(q_n)\rangle$). We denote finite binary sequences by σ, τ. In this context, $|\sigma|$ denotes the length of the σ. If σ is an initial segment of τ (or $B \in 2^\omega$), we denote this by $\sigma \prec \tau$ ($\sigma \prec B$). We will use the following to denote intervals in 2^ω: $I(\sigma) = \{B \in 2^\omega \mid \sigma \prec B\}$.

$\{\phi_0, \phi_1, \dots, \phi_n, \dots\}$ enumerates the the partial recursive functions, and ϕ denotes an arbitrary partial recursive function. $\{\Phi_0, \Phi_1, \dots, \Phi_n, \dots\}$ enumerates the corresponding functions on $I_{\mathbb{Q}}$, with Φ denoting an arbitrary partial recursive function with domain and range contained in $I_{\mathbb{Q}}$. M (with adornments) will denote a (standard) recursive Turing machine, thought of as a machine which infers a class

of functions. Such machines M will usually be used in the context of the standard inference classes, operating on sequences of natural numbers, and producing natural numbers as outputs. We will also use them in the context of standard inference of functions on $I_{\mathbb{Q}}$. We will often use the notation "$\overset{\infty}{\forall} x$" in this context to mean "for all but finitely many x". G (with adornments) denotes a recursive Turing machine with inputs from $[I_{\mathbb{Q}}]^{<\omega}$ and outputs in $I_{\mathbb{Q}}$ or $\mathbb{N}$, depending on whether its guesses are to be function values, or indices of functions.

REC will denote the standard (total) recursive functions (i.e. those recursive on $\mathbb{N}$). $\mathbb{Q}REC$ will denote the class of total recursive functions on $I_{\mathbb{Q}}$; we will also call these the *recursive rational-valued functions.* We shall denote the subclass of recursive $\{0,1\}$-valued functions by $\mathbb{Q}SET$. RC denotes the recursive rational functions which are *continuous* on $I_{\mathbb{Q}}$, and RUC denotes those which are *uniformly continuous* on $I_{\mathbb{Q}}$. Since $I_{\mathbb{Q}}$ is neither compact nor connected, continuity does not imply uniform continuity, so that $RUC \subsetneq RC$. Note that we may embed REC into RUC by mapping f to $f_{\mathbb{Q}}$, and then mapping $f_{\mathbb{Q}}$ to $\hat{f}_{\mathbb{Q}}$ defined by $\hat{f}_{\mathbb{Q}}(0) = 0$, and for $n > 0$, $\hat{f}_{\mathbb{Q}}(1/n) = f_{\mathbb{Q}}(q_{n-1})/n$, linearly interpolating these values for other rationals.

See Odifreddi [16], Rogers [21] or Soare [24] for developments of basic recursion theory. All notions from measure theory, real analysis, and topology can be found in Halmos [11], Royden [20], and Munkres [15], respectively.

2.2 Inductive Inference

We give a brief summary of the standard inductive inference classes and the theorems relating them. The three basic notions of inductive inference are as follows:

DEFINITION 2.2.1. A class $\mathcal{C}$ of recursive functions is *next-value identifiable* ($\mathcal{C} \in NV$) if there is an $M : [\omega]^{<\omega} \longrightarrow \omega$ such that for every $f \in \mathcal{C}$,

$$\overset{\infty}{\forall} n \; f(n) = M(\langle f(0), \ldots, f(n-1)\rangle).$$

Thus, an NV-inference machine M tries to correctly guess the sequence $\{f(0), f(1), \dots\}$, making only finitely many errors. We also say that M NV-identifies $\mathcal{C}$, and that $\mathcal{C} \in NV$ via M.

DEFINITION 2.2.2. A class $\mathcal{C}$ of recursive functions is *explanatorily identifiable* ($\mathcal{C} \in EX$) if there is a $M : [\omega]^{<\omega} \longrightarrow \omega$ so that for every $f \in \mathcal{C}$, there is an index e of f such that

$$\overset{\infty}{\forall} n \; M(\langle f(0), \dots, f(n-1)\rangle) = e.$$

Thus, an EX-inference machine M tries to settle on a correct index for the input f, making only finitely many errors. Gold [10] also introduces the notion of explanatory consistency:

DEFINITION 2.2.3. A class $\mathcal{C}$ of recursive functions is *identifiable by consistent explanation* ($\mathcal{C} \in EX_{cons}$) if $\mathcal{C} \in EX$ via M, and for all n,

$$\phi_{M(\langle f(0),\dots,f(n-1)\rangle)}|_n = f|_n.$$

Feldman [8] further weakens the notion of EX-inference as follows:

DEFINITION 2.2.4. A class $\mathcal{C}$ of recursive functions is *behaviorally correctly identifiable* ($\mathcal{C} \in BC$) if there is a $M : [\omega]^{<\omega} \longrightarrow \omega$ so that for every $f \in \mathcal{C}$,

$$\overset{\infty}{\forall} n \; \phi_{M(\langle f(0),\dots,f(n-1)\rangle)} = f.$$

Thus, a BC-inference machine M must, after finitely many stages, output only indices of the input function f. However, unlike the EX case, we may output *any* index of f, infinitely often changing our mind about which index to output.

The notions PEX and PBC are similar to EX and BC, except that M may only output indices of *total* functions. Theorems of Blum and Blum [4], Barzdin [3], Case and Smith [6], Gold [10], and Harrington [12] yield the following hierarchy of inference classes:

PROPOSITION 2.2.5. $NV = PEX \subsetneq EX_{cons} \subsetneq EX \subsetneq BC.$

Blum and Blum [4], and Case and Smith [6] introduce the concept of inference with anomalies:

DEFINITION 2.2.6. If f is total, we say that ϕ_e is an *n-variant* of f ($\phi_e \simeq^n f$) if

$$|\{x \mid \phi_e(x)\uparrow \text{ or } \phi_e(x)\downarrow\neq f(x)\}| \leq n.$$

We say simply that ϕ_e is a *variant* of f ($\phi_e \simeq^* f$) if ϕ_e is an n-variant of f for some n.

DEFINITION 2.2.7. A class $\mathcal{C}$ of recursive functions is *explanatorily identifiable with n errors* ($\mathcal{C} \in EX^n$) if there is an $M : [\omega]^{<\omega} \longrightarrow \omega$ so that for every $f \in \mathcal{C}$, there is an index e such that $\phi_e \simeq^n f$, and

$$\overset{\infty}{\forall} n\ M(\langle f(0), \dots, f(n-1)\rangle) = e.$$

DEFINITION 2.2.8. A class $\mathcal{C}$ of recursive functions is *explanatorily identifiable with finitely many errors* ($\mathcal{C} \in EX^*$) if there is an $M : [\omega]^{<\omega} \longrightarrow \omega$ so that for every $f \in \mathcal{C}$, there is an index e such that $\phi_e \simeq^* f$, and

$$\overset{\infty}{\forall} n\ M(\langle f(0), \dots, f(n-1)\rangle) = e.$$

Note that $EX^0 = EX$. The aforementioned authors have shown the following:

PROPOSITION 2.2.9. For all n,

- $EX^n \subsetneq EX^*$ (in fact, $\cup_n EX^n \subsetneq EX^*$), and
- $EX^n \subsetneq EX^{n+1}$.

Case and Smith [6] show that $EX^* \subsetneq BC$. They also extend the concept of inference with anomalies to BC — BC^n and BC^* are defined similarly to their EX-counterparts. The previous proposition also holds for BC^n and BC^*. Harrington [12] shows that BC^* is as weak a notion of inference as possible, in the sense that $REC \in BC^*$.

We thus have the following hierarchy of inference classes:

$$
\begin{aligned}
NV = PEX &\subsetneq EX_{cons} \\
&\subsetneq EX \subsetneq EX^1 \subsetneq EX^2 \subsetneq \ldots \subsetneq \cup_n EX^n \subsetneq EX^* \\
&\subsetneq BC \subsetneq BC^1 \subsetneq BC^2 \subsetneq \ldots \subsetneq \cup_n BC^n \subsetneq BC^*.
\end{aligned}
$$

In the sequel, we shall see that the notions of approximate inference yield a hierarchy which is no longer linear.

2.3 Inference with Oracles

Fortnow et al. [9], and Kummer and Stephan [14] exanmine the concept of oracle inductive inference, in which the inference machine M is allowed queries to an arbitrary set $A \subset \mathbb{N}$.

DEFINITION 2.3.1. We say that an oracle Turing machine (O.T.M.) $M^{()}$ is *categorical* if for every $A \subset \mathbb{N}$, M^A is total (in the sequel, we shall usually identify such A with its characteristic function).

We give the definitions of $NV[A]$, $NV[A*]$ here. The definitions for $EX[A]$, $EX[A*]$, $BC[A]$, and $BC[A*]$ are obtained from the definitions of EX and BC similarly.

DEFINITION 2.3.2. Let $A \subset \mathbb{N}$. A class $\mathcal{C}$ of recursive functions is *next-value identifiable from A* ($\mathcal{C} \in NV[A]$) if there is a categorical O.T.M. $M^{()}$ so that for every $f \in \mathcal{C}$,

$$\overset{\infty}{\forall} n \; f(n) = M^A(\langle f(0), \ldots, f(n-1) \rangle).$$

DEFINITION 2.3.3. Let $A \subset \mathbb{N}$. A class $\mathcal{C}$ of recursive functions is *next-value identifiable from A with finitely many queries* ($\mathcal{C} \in NV[A*]$) if there is a categorical O.T.M. $M^{()}$ so that $\mathcal{C} \in NV[A]$ via M^A, and for each $f \in \mathcal{C}$, there is an n so that f is inferred by $M^{A \cap \{0,\ldots,n-1\}}$.

For oracle inference machines, the fundamental task is to determine which oracles yield no increase in inference power and which ones allow us to infer all of *REC*.

DEFINITION 2.3.4. We say that A is *NV-trivial* if $NV[A] = NV$, and *NV-omniscient* if $REC \in NV[A]$. We define these notions for other inference criteria similarly.

K denotes the Halting problem, $K = \{\langle x, y \rangle \mid \phi_x(y)\downarrow\}$. It is well-known that K is EX-omniscient. An O.T.M. which EX-identifies REC using K may be defined as follows (on input $f \in REC$):

Stage n:

Case 1: There is an $e < n$ such that $\phi_e(y) \downarrow= f(y)$ for all $y < n$ (use K to determine this). Output the least such.

Case 2: No such e exists. Output n.

End of Construction.

As another example, consider

$$TOT = \{e \mid \forall y\, \phi_e(y)\downarrow\}.$$

It is easy to see that $REC \in EX[TOT*]$. The machine to infer REC is as follows (input $f \in REC$, and let $e_0 = 0$):

Stage n:

Case 1: $e_0 \in TOT$, and for all $y < n$ $\phi_{e_0}(y) = f(y)$. Output e_0.

Case 2: There is an $e < n$ such that $e \in TOT$, and for all $y < n$ $\phi_e(y) = f(y)$. Output the least such e, and set $e_0 = e$.

Case 3: No such e exists. Output e_0.

End of Construction.

DEFINITION 2.3.5. $G \in 2^\omega$ is *generic* if for each Σ_1 $W \subset \{0,1\}^*$, either

- $(\exists \sigma \prec G)\, (\forall \tau \succeq \sigma)\, \tau \notin W$, or
- $(\exists \sigma \prec G)\, \sigma \in W$.

Following Fortnow et al. [9], we use use the notation $\mathcal{G}(A)$ to mean that either A is recursive, or $A \leq_T K$ and $A \equiv_T G$ for some generic G.

The main result in Fortnow et al. [9] is that $EX[A] = EX \Leftrightarrow \mathcal{G}(A)$ and $BC[A] = BC \Leftrightarrow \mathcal{G}(A)$. The crux of the argument is to show that for generic G, $EX[G*] = EX[G]$ and that $BC[G*] = BC[G]$.

CHAPTER 3
INFERENCE OF RECURSIVELY CONTINUOUS FUNCTIONS

3.1 The Recursively Continuous Functions

We recall the definition of a recursively continuous function on the rational unit interval $I_{\mathbb{Q}}$:

DEFINITION 3.1.1. f is *recursively continuous* if f is recursive, and uniformly continuous w.r.t. the standard metric on $\mathbb{Q}$. We denote the class of all such functions by *RUC*.

There are other versions of continuity which we could consider in place of uniform continuity in the above definition. Note that any uniformly continuous function on $I_{\mathbb{Q}}$ has a unique uniformly continuous extension to $[0, 1]$. However, an f which is continuous on $\mathbb{Q}$, but not uniformly so, may have no continuous extension to all of $[0, 1]$. Thus, with our present definition, we may identify $f \in RUC$ with its unique continuous extension to $[0, 1]$. We could also consider a more standard definition of a "recursively continuous" real function. For example, Cenzer and Remmel[7] define a real-valued f to be recursively continuous if there is a uniformly recursive sequence $\{f_n\}$ of functions on $\mathbb{Q}$, and a recursive modulus of convergence e such that for any $r \in \mathbb{R}$, and $q \in \mathbb{Q}$, if $m > e(n)$, then

$$|q - r| < \frac{1}{e(n)} \Rightarrow |f_m(q) - f(r)| < \frac{1}{n}.$$

Using this definition, for a *rational* input r, we could then approximate $f(r)$ to within any desired tolerance. However, this does not allow us to directly compute $f(r)$, even if $f(r)$ is rational as well. Furthermore, functions which are recursively continuous in this sense are not necessarily recursive in the usual sense, and in particular, may not be in the class *RUC* defined above. Since we will be considering only recursive

functions in the sequel, we allow our definition of RUC to stand. See also Pour-El and Richards [18] for alternate definitions of recursive continuity.

In the experimental sciences, and in many engineering applications, the elements of a class of phenomena under consideration are assumed to be continuous. Thus, it is natural to consider the effect of restricting the usual inference notions to such a class. In this chapter, we will show that RUC is a domain in which all of the standard inference notions can be formulated and in which the standard inference hierarchy given in the last section is preserved. This is in some sense a measure of the nontriviality of RUC: if, for example, we were to restrict inductive inference to the class of polynomials over $\mathbb{N}$ or to the primitive recursive functions, the inference hierarchy would collapse. We first show some basic properties of RUC.

Note that $RUC \neq \mathbb{Q}REC$: if $A \subset \omega$ is finite, then χ_A is recursive, but not continuous. The class RUC is Π^0_3:

$$e \in RUC \Leftrightarrow e \in TOT \ \ \& \ \forall m \exists n \forall k, l$$

$$|q_k - q_l| < 1/n \Rightarrow |\phi_e(q_k) - \phi_e(q_l)| < 1/m.$$

In fact, RUC is Π^0_3-complete, as we will show in the next theorem.

THEOREM 3.1.2. The class RUC is Π^0_3-complete.

PROOF. We first construct a computable irrational r and a computable sequence of irrationals $\{r_n\}_{n\in\omega}$ s.t. $0 < r_0 < r_2 < \cdots < r$, $1 > r_1 > r_3 > \cdots > r$, with $r_{2n} \to r$ and $r_{2n+1} \to r$. Let $r = .1010010001000\cdots$ (clearly, r is computable). Now, define the sequence $\{r_k\}_{k\in\omega}$ by letting r_{2k} be the irrational formed by replacing the k^{th} 1 in the binary expansion of r by 0, and letting r_{2k+1} be the irrational formed by replacing the k^{th} 0 in the binary expansion of r by 1. Then $\{r_k\}_{k\in\omega}$ is clearly computable. Let $E_0 = [0, r_0)$, $E_{2k} = [r_{2k-2}, r_{2k})$, $E_1 = (r_1, 1]$, and $E_{2k+1} = (r_{2k+1}, r_{2k-1}]$.

Given a Π^0_3 set A, we shall, uniformly in a, construct a recursive function f_a s.t. $f_a \in RUC$ iff $a \in A$. The essence of the construction is as follows. If A is Π^0_3, then there is a recursive relation R s.t. $a \in A \Leftrightarrow \forall i \exists j \forall k \ R(i, j, k, a)$. We will define

a recursive map $f_a : I_{\mathbb{Q}} \longrightarrow I_{\mathbb{Q}}$ which is uniformly continuous if and only if $a \in A$. On the interval E_i, f_a consists of a series of "sawtooths" of height $1/2^i$, possibly interspersed with "gaps" (sub-intervals on which f_a is identically zero). To make f_a recursive, we define f_a on each rational q_k at no later than stage k of the construction. Now, if q is the first rational enumerated into E_i, we let $f_a \equiv 0$ on $E_i \cap \mathbb{Q}$. Initially, we will try to satisfy the relation R in E_i with $j = 0$. Since j will depend on i, we will denote j in the i^{th} interval by j_i. A "sawtooth" is added in the interval E_i whenever we find a new witness that $a \notin A$, that is, a new $q = q_k$ enumerated into E_i such that $(\exists \hat{k} < k)\neg R(i, j_i, \hat{k}, a)$. In this case, if f_a is not already defined at q_k, f_a is extended by a "sawtooth" in E_i from the greatest rational previously enumerated into the interval to the new point q_k, and j_i is incremented (since the current j_i can no longer witness that $a \in A$). If no such witness is found, i.e. if $(\forall \hat{k} < k)R(i, j_i, \hat{k}, a)$, a "gap" is added — in other words, f_a is extended by the zero function to q_k.

Now, if $a \notin A$, then there will be at least one i_0 such that to each j there will correspond a k_j such that $\neg R(i_0, j, k_j, a)$. Since $E_{i_0} \cap \mathbb{Q}$ is dense in E_{i_0}, for any j there will be a q_k enumerated into E_{i_0} with $k_j < k$, and with f_a as yet undefined at q_k. Thus, f_a will have an infinite number of "sawtooths" of the same height on E_{i_0}. But then $\lim_{q \to r_{i_0}}$ does not exist, so that f_a will is not uniformly continuous, hence not recursively continuous. On the other hand, if $a \in A$, for each i, there will be a stage N_i in our construction after which j_i is no longer incremented (j_i will be the least j for which $R(i, j, k, a)$ holds for all k). Thus, for $q = q_k$ enumerated into E_i after stage N_i, $(\forall \hat{k} < k)R(i, j_i, \hat{k}, a)$ will hold. So, if $a \in A$, f_a will be zero for $q > q_{N_i}$ in each E_i. Thus, f_a is continuous at every rational in E_i, and

- $\lim_{q \to r_i} = 0$ at each r_i, and
- $\lim_{q \to r} = 0$, since the "sawtooths" in each E_i decrease in height to zero as $r_i \to r$.

We now give a more formal algorithm for f_a. The function s in the algorithm will provide the precise "sawtooth" needed at stage k. We define s by

$$s(x) = \begin{cases} 2x, & \text{if } 0 \leq x \leq 1/2, \\ -2x+2, & \text{if } 1/2 < x \leq 1, \end{cases}$$

and scale s to suit our needs at any particular stage.

Stage k: First, we must determine to which E_i the rational q_k belongs. This is effective, since the $\{r_i\}$ are uniformly computable. Once this is determined, we determine $f_a(q_k)$ as follows:

Case 1: q_k is the first rational enumerated into E_i. Then set $f_a(x) = 0$ for all $x \leq q_k$ in E_i. Set $x_i = q_k$, and set $j_i = 0$.

Case 2: $q_k \leq x_i$. In this case, do nothing, since $f_a(q_k)$ is defined at a previous stage.

Case 3: $q_k > x_i$ and $(\forall \hat{k} < k)\ R(i, j_i, k, a)$. Then set $f_a(x) = 0$ for all x in the interval $[x_i, q_k]$. Set $x_i = q_k$.

Case 4: $q_k > x_i$ and $(\exists \hat{k} < k)\ \neg R(i, j_i, k, a)$. Then extend f_a on the interval $[x_i, q_k]$ using the "sawtooth" function:

$$\frac{1}{2^i} s\left(\frac{x - x_i}{q_k - x_i}\right).$$

Set $x_i = q_k$, and increment j_i.

The function f_a is uniformly recursive in a, so we have defined a binary recursive function $\psi(a, n)$. Thus, by the S^m_n Theorem, there is a recursive f s.t. $\phi_{f(a)}(n) = \psi(a, n)$. But then, f is the desired reducibility, for if $a \in A$, then for each i, the last case above is never satisfied beyond some stage k_i. Thus, $\phi_{f(a)}$ is eventually zero in each interval, so $\phi_{f(a)}$ is continuous on $I_{\mathbb{Q}}$. On the other hand, if $a \notin A$, then there is an i s.t. the last case comes up infinitely often, and so $\lim_{x \to r_i} \phi_{f(a)}(x)$ does not exist, whence $\phi_{f(a)}$ is not continuous (see figure 3.1). $\square$

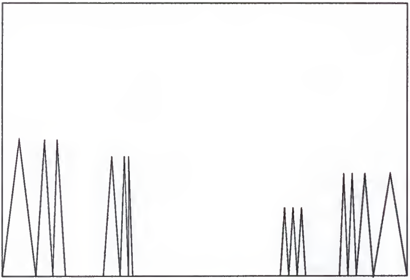

Figure 3.1: The construction of f.

In the sequel, it will be useful to encode a recursive function into a recursively continuous one. In fact, we shall produce an effective embedding ε of REC into RUC, allowing us to perform the encoding in a uniform way. Furthermore, it will be easy to decode the values of the recursive function from the RUC function into which it has been embedded. First, note that we may effectively obtain a strictly decreasing sequence $\{\hat{q}_j\}_{j\in\omega}$ from our enumeration $\{q_i\}_{i\in\omega}$ of $I_{\mathbb{Q}}$, such that $\hat{q}_0 = 1$ and $\lim_{j\to\infty}\hat{q}_j = 0$. Now, given $f \in REC$, we define $\hat{f} = \varepsilon(f)$ on the $\{\hat{q}_j\}_{j\in\omega}$ by

$$\hat{f}(\hat{q}_j) = \frac{q_{f(j)}}{2^j},$$

and by linear interpolation on rationals between these points. Finally, we set $\hat{f}(0) = 0$. Clearly, the map ε is one-to-one, and for any $f \in REC$, $\hat{f}$ is recursively continuous. ε also has an effective inverse, since $q_{f(j)} = 2^j \hat{f}(\hat{q}_j)$. We next consider the effect of restricting the standard notions of inference to the class RUC.

3.2 Next Value Inference of *RUC*

Let us consider the inference class NV^{RUC}, which is simply the class *NV*, restricted to the recursively continuous functions. As mentioned in the Introduction, $REC \notin NV$. We now prove the analog for NV^{RUC}.

THEOREM 3.2.1. $RUC \notin NV^{RUC}$.

PROOF. Given a recursive M, we will construct a recursively continuous f for which M is not a next-value function. We will give the value of f on q_{n+1} by linear interpolation, using the points $(q_0, f(q_0)), (q_1, f(q_1)), \ldots, (q_n, f(q_n))$. First, let $f(0) = 0$, $f(1) = 1$. Now, suppose that $f(q_0), f(q_1), \ldots, f(q_n)$ have been determined, and that the set $\{q_0, q_1, \ldots, q_n\}$ is ordered as $0 = q_{k_0} < q_{k_1} < \ldots < q_{k_n}$. Let $h_{n,i}(x)$ denote the linear interpolation of the points $(q_{n_i}, f(q_{n_i}))$ and $(q_{n_{i+1}}, f(q_{n_{i+1}}))$, and h_n the function obtained by "pasting" the $h_{n,i}$ together. If $M(\langle f(q_0), \ldots, f(q_n)\rangle)$ does not equal $h_n(q_{n+1})$, set $f(q_{n+1}) = h_n(q_{n+1})$. If, however, $M(\langle f(q_0), \ldots, f(q_n)\rangle) = h_n(q_{n+1})$, set $f(q_{n+1} = h_n(q_{n+1}) + \frac{1}{2^{n+1}}$ if $h_n(q_{n+1}) + \frac{1}{2^{n+1}} \leq 1$, otherwise, set $f(q_{n+1}) = h_n(q_{n+1}) - \frac{1}{2^{n+1}}$.

The function f is recursive, since $f(q_n)$ is defined at stage n. Note that $|h_n(x) - h_{n+1}(x)| < 2^{-n-1}$ for each n and x, so that the real-valued, continuous functions $\{h_n\}$ converge uniformly to a real-valued, continuous function h on $[0,1]$. Then h is uniformly continuous on $[0,1]$. Thus, since f is just the restriction of h to $I_{\mathbb{Q}}$, f is uniformly continuous on $I_{\mathbb{Q}}$, i.e. $f \in RUC$. By construction, M is not a next-value function for f. Since M was arbitrary, we have shown that no recursive Turing machine can be a next-value function for all of *RUC*, which proves the theorem. □

3.3 Inference of *RUC* by Consistent Explanation

We now consider EX_{cons}^{RUC}, the restriction of EX_{cons} to *RUC*. The proof that $RUC \notin EX_{cons}^{RUC}$ uses a method similar to the one used in the proof by Odifreddi [17] that $REC \notin EX_{cons}$.

THEOREM 3.3.1. $RUC \notin EX_{cons}^{RUC}$.

PROOF. We will assume that M witnesses that $RUC \in EX_{cons}^{RUC}$ and obtain a contradiction by constructing a recursively continuous function f in a manner which prohibits M from stabilizing on any index for f. We will accomplish this by constructing a sequence $\{q_{n_k}\}$ so that $M(\langle f(q_0), \dots, f(q_{n_k})\rangle) \neq M(\langle f(q_0), \dots, f(q_{q_{n_{k+1}}})\rangle)$, thus ensuring that

$$\lim_{n\to\infty} M(\langle f(q_0), \dots, f(q_n)\rangle)$$

does not exist. As we have done before, we will construct a sequence of real-valued functions $\{h_n\}$ such that the limiting function h is continuous, with f being the restriction of h to the rationals. h_0 will be identically zero. At stage n we extend h_n past q_n in two different ways, with $\hat{h}$ denoting the extension which agrees with h_n, and h' denoting a function obtained by perturbing h_n past q_n by a "hat" function of height 2^{-n}. The restrictions to $\mathbb{Q}$ of $\hat{h}$ and h' will both be recursively continuous, and different, so M must eventually produce different indices for them. We can then effectively find $n_k =$ the least such n such that M produces different indices. We then set h_{n+1} equal to whichever of $\hat{h}, h'$ yielded an index different from the index produced by M at stage n_{k-1}, and let $f(q_n) = h_{n+1}(q_n)$. Thus f will be recursive, and uniformly continuous, but M will not identify it, showing that $RUC \notin EX_{cons}^{RUC}$.

We now give the details of the construction. Let

$$\hat{q}_k = \max\{q_0, q_1, \dots, q_k\} \setminus \{1\}.$$

At stage k in the construction, there are two possible cases. If $f(q_k)$ is already defined, do nothing. On the other hand, if $f(q_k)$ is not yet defined, we extend f continuously past $\hat{q}_{k-1}$ in two different ways, by $\hat{h}$ and h' as described above. For notational convenience, set $h_{-1} \equiv 0$, $n_{-1} = 0$, and let s denote the "sawtooth" function of Theorem 3.1.2. Also, let $e_k = M(\langle f(q_0), \dots, f(q_{k-1})\rangle)$. We begin our construction with $f(0) = f(1) = 0$.

Define the functions $h_k, \hat{h}, h'$ at each stage k such that $f(q_k)$ is undefined as follows (at stages k where $f(q_k)$ is already defined, do nothing):

$$\begin{aligned}\hat{h}(x) &= \begin{cases} h_{k-1}(x), & \text{if } \hat{q}_k = \hat{q}_{k-1} \text{ or } x \leq \hat{q}_k, \\ \frac{1}{2^k} s\left(\frac{x-\hat{q}_k}{1-\hat{q}_k}\right), & \text{otherwise.}\end{cases} \\ h'(x) &= h_{k-1}(x)\end{aligned}$$

Now, we can effectively find $n_k > n_{k-1}$ such that

$$a_1 = M(\langle \hat{h}(q_0), \ldots, \hat{h}(q_{n_k})\rangle) \neq M(\langle h'(q_0), \ldots, h'(q_{n_k})\rangle) = a_2.$$

Let $h_k = \hat{h}$ if $a_1 \neq e_k$, and h' otherwise. Let $f(q_k) = h_k(q_k)$, thus f is recursive. As in the previous construction, we see that $|h_k(x) - h_{k+1}(x)| < 2^{-k-1}$ for each k. Thus, if we set $h(x) = \lim_{k\to\infty} h_k(x)$, then h is the uniform limit of continuous functions, so h is continuous, hence uniformly continuous on $[0,1]$. But f is just the restriction of h to $I_{\mathbb{Q}}$, so that f is also uniformly continuous on $I_{\mathbb{Q}}$, whence $f \in RUC$. By construction, M does not identify f, which proves the theorem. □

As noted previously, if we modify the definition of EX_{cons} so that the inference machine M must always output the index of a total recursive function (calling the new inference notion PEX_{cons}), we have $PEX_{cons} = NV$. Restricting to RUC, it is easy to show that $PEX_{cons}^{RUC} = NV^{RUC}$.

THEOREM 3.3.2. $PEX_{cons}^{RUC} = NV^{RUC}$.

PROOF. Let M NV^{RUC}-identify $\mathcal{C}$. For any sequence $a = \langle a_0, \ldots, a_{n-1}\rangle$, let $h(a)$ be the least $i < n$ s.t. $\forall i \leq j < n$ $M(\langle a_0, \ldots, a_{j-1}\rangle) = a_j$, if one exists; otherwise let $h(a) = n$. Let

$$\psi(a,x) = \begin{cases} a_x, & \text{if } a \text{ is a sequence of length } n \text{ and } x < h(a), \\ M(a), & \text{if } a \text{ is a sequence number, and } x \geq h(a), \\ 0, & \text{otherwise.}\end{cases}$$

By the S_n^m theorem, there is a recursive $\hat{M}$ s.t. $\phi_{\hat{M}(e)}(x) \simeq \psi(e,x)$. Note that $\hat{M}(n)$ is a recursive index for all n, since ψ is a recursive function of two variables. Furthermore, if $f \in C$, there is an N s.t. M agrees with f for $n > N$. If N is minimal,

then $h(\langle f(0), f(1), \ldots, f(n)\rangle) = N$ for all $n > N$, so that $\hat{M}(\langle f(0), \ldots, f(n)\rangle)$ is an index for f at each stage $n > N$. Thus, we need only to ensure that these indices stabilize in order that $\mathcal{C} \in PEX_{cons}^{RUC}$. To accomplish this, at stage n we simply output the first $e = \langle f(0), \ldots, f(k)\rangle \leq n$ such that f agrees with $\phi_{\hat{M}(e)}$ up to n.

In the other direction, if M PEX_{cons}^{RUC}-identifies $\mathcal{C}$, then at stage $n+1$ we output $\phi_{M(f(q_0), f(q_1), \ldots, f(q_n))}(n+1)$. This is recursive, since $M(f(q_0), f(q_1), \ldots, f(q_n)) \in TOT$, and it NV^{RUC}-identifies $\mathcal{C}$, since if $f \in \mathcal{C}$, then M stabilizes on an index of f after some stage N, so that for all $n > N$ $\phi_{M(f(q_0), f(q_1), \ldots, f(q_N))}(n+1) = f(n+1)$. □

Recall that $NV \subsetneq EX_{cons}$. We note that the analogous inclusion $NV^{RUC} \subsetneq EX_{cons}^{RUC}$ for the relativized inference types holds as well. The proof that $NV \subsetneq EX_{cons}$ given by Odifreddi [17] does not easily relativize. Our proof will make use of the encoding function $\varepsilon : REC \longrightarrow RUC$ given above. For $\mathcal{C} \subseteq REC$, let $\hat{\mathcal{C}}$ denote the set $\{\varepsilon(f) | f \in \mathcal{C}\}$. Then $NV^{RUC} \subsetneq EX_{cons}^{RUC}$ is an easy corollary of the lemmas that follow.

DEFINITION 3.3.3. For any M, $\sigma = \langle a_0, \ldots, a_{m-1}\rangle$, we define $S_{M,\sigma}$ on REC by recursion as follows

$$S_{M,\sigma}(n) = \begin{cases} M(\langle a_0, \ldots, a_{n-1}\rangle), & \text{if } n < m, \\ M(\langle a_0, \ldots, a_{m-1}, S_{M,\sigma}(m), \ldots, \\ \quad S_{M,\sigma}(n-1)\rangle), & \text{otherwise.} \end{cases}$$

LEMMA 3.3.4. $\mathcal{C} \in NV \Leftrightarrow \hat{\mathcal{C}} \in NV^{RUC}$.

PROOF. Suppose M NV-identifies $\mathcal{C}$, and let $h \in \hat{\mathcal{C}}$. At stage $n+1$, we are given the values $h(q_0), h(q_1), \ldots, h(q_n)$ from which to make a prediction for $h(q_{n+1})$. Among these values will be $h(\hat{q}_0), h(\hat{q}_1), \ldots, h(\hat{q}_{k_n})$ for some k_n. But $h = \varepsilon(f)$ for some $f \in \mathcal{C}$, so we may effectively find $f(i)$ from $h(\hat{q}_i)$, using the construction of ε. Now, let $f_n = S_{M,\langle f(0), \ldots, f(k_n)\rangle}$. Then, as a guess for $h(q_{n+1})$ at this stage, output $\{\varepsilon(f_n)\}(q_{n+1})$. Since M is eventually correct on input $f(0), f(1), \ldots, f(n)$, this procedure is eventually correct on input $h(q_0), h(q_1), \ldots, h(q_n)$.

Similarly, if $\hat{\mathcal{C}} \in NV^{RUC}$ via G, we can construct an M to NV-identify $\mathcal{C}$. □

LEMMA 3.3.5. $\mathcal{C} \in EX_{cons} \Leftrightarrow \hat{\mathcal{C}} \in EX_{cons}^{RUC}$.

PROOF. Suppose M EX_{cons}-identifies M, and let $h \in \hat{\mathcal{C}}$. At stage $n+1$ we wish to produce an index for h. The procedure is similar to the one given in the previous lemma. Now, $h = \varepsilon(f)$ for some $f \in \mathcal{C}$, so from $h(q_0), h(q_1), \ldots, h(q_n)$ we may effectively produce $f(0), f(1), \ldots, f(n)$. We then apply M to the sequence of f-values to produce a guess e_n for an index of f. Using some fixed index of ε, viewed as a functional on the partial recursive functions, we then produce an index $\hat{e_n}$ from e_n. Eventually, the e_n's will converge to an index e for f, so that $\hat{e}$ is an index of h. The other direction is shown similarly. □

THEOREM 3.3.6. $EX_{cons}^{RUC} - NV^{RUC} \neq \emptyset$.

PROOF. By way of contradiction, suppose that $EX_{cons}^{RUC} = NV^{RUC}$. Then by the previous theorem, $EX_{cons}^{RUC} = PEX_{cons}^{RUC}$, so that by the two lemmas above,

$$\mathcal{C} \in NV \iff \hat{\mathcal{C}} \in NV^{RUC} \iff \hat{\mathcal{C}} \in EX_{cons}^{RUC} \iff \mathcal{C} \in EX_{cons}.$$

But then $NV = EX_{cons}$, a contradiction. □

The inclusions among the various inference types defined in this chapter can be summarized as follows:

$$PEX_{cons}^{RUC} = NV^{RUC} \subsetneq EX_{cons}^{RUC}.$$

Now, we may define the classes EX, EX^n, EX^*, BC, BC^n, and BC^* relative to RUC in the usual way. Then, using the procedure in Lemma 3.3.5, it is easy to see that the inclusions among the various classes relativize also. In the next chapter, we will formulate new notions of inference which allow us to infer RUC.

CHAPTER 4
APPROXIMATE INFERENCE OF RECURSIVE FUNCTIONS

4.1 Notions of Approximate Inference

We are now in a position to define notions of inference, NV_∞, EX_∞, and BC_∞, which are in a sense "continuous" analogues of the standard notions *NV*, *EX*, and *BC*. The idea for these new notions comes from the following situation. Suppose that $f : [0,1] \longrightarrow [0,1]$ is continuous, and we are allowed to ask "What is $f(x)$?" for (only) countably many x. We may then ask, in this manner, for $f(0)$, $f(1)$, $f(1/2)$, $f(1/3)$, $f(2/3)$, and so on. After we ask for the n^{th} value, we can then form an approximation f_n to the graph of f, perhaps by splines, or by a polynomial interpolation. Since we know that f is continuous, we know that the sequence of continuous functions $\{f_n\}$ converges uniformly to f. Thus, we have a procedure whereby we can build reasonable approximations to f in stages, with the knowledge that, "in the limit", we recover f itself. With this procedure in mind, we define the ***approximate*** inference classes.

DEFINITION 4.1.1. A class $\mathcal{C}$ of recursive rational functions is ***next-value approximable*** ($\mathcal{C} \in NV_\infty$) if there is a $G : [I_{\mathbb{Q}}]^{<\omega} \longrightarrow I_{\mathbb{Q}}$ such that for every $f \in \mathcal{C}$

$$\lim_{n\to\infty} |f(q_n) - G(\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle)| = 0.$$

This definition compares to standard NV-inference, as applied to rational functions, where G infers f only if

$$f(q_n) = G(\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle)$$

for all but finitely many n.

DEFINITION 4.1.2. A class $\mathcal{C}$ of recursive rational functions is *explanatorily approximable* ($\mathcal{C} \in EX_\infty$) if there is a recursive $G : [I_\mathbb{Q}]^{<\omega} \longrightarrow \mathbb{N}$ so that for every $f \in \mathcal{C}$, there is an index e such that

- Φ_e is total,
- for all but finitely many n, $G(\langle f(q_0), \ldots , f(q_n)\rangle) = e$, and
- $\lim_{n\to\infty} |f(q_n) - \Phi_e(q_n)| = 0$.

Thus, an EX_∞-inference machine behaves like an EX-inference machine, except that the index it settles on must only approximate the input f in the NV_∞ sense. The definition for $\mathcal{C} \in PEX_\infty$ is similar, but G must output only indices for *total* functions.

DEFINITION 4.1.3. A class $\mathcal{C}$ of recursive rational functions is *behaviorally correctly approximable* ($\mathcal{C} \in BC_\infty$) if there is a recursive $G : [I_\mathbb{Q}]^{<\omega} \longrightarrow \mathbb{N}$ such that for every $f \in \mathcal{C}$,

$$\lim_{n\to\infty} \left\|f - \Phi_{G(\langle f(q_0), f(q_1), \ldots, f(q_n)\rangle)}\right\|_\infty = 0,$$

where $\|\Phi\|_\infty = \sup_{x\in I_\mathbb{Q}}\{|\Phi(x)|\}$ if Φ is total, and $\|\Phi\|_\infty = 1$ if Φ is partial. The definition for $\mathcal{C} \in PBC_\infty$ is similar, but G must output only indices for *total* functions.

This definition compares to standard BC-inference, as applied to rational functions, where G infers f only if

$$f = \Phi_{G(\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle)}$$

for all but finitely many n. We recall that the more restrictive notion of EX-inference requires further that there is a fixed e such that $G(\langle f(q_0), f(q_1), \ldots , f(q_{n-1})\rangle) = e$ for all but finitely many n. The less restrictive notion of BC^*-inference allows G to infer f if, for all but finitely many n,

$$f(x) = \Phi_{G(\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle)}(x)$$

except on a finite set. We note again that the class $\mathbb{Q}REC$ is in BC^*, but that neither $\mathbb{Q}REC$ nor $\mathbb{Q}SET$ are in BC.

An obvious way to weaken the above definition is to require only that a BC_∞ machine G output guesses $\{f_n\}$ so that the f_n's converge *pointwise* to the input function f. The modified inference notion is then equivalent to BC^*, for if $f \in \mathbb{Q}REC$, we approximate f at stage n by f_n, defined by

$$f_n(q_k) \;=\; \begin{cases} f(q_k) & k < n, \\ 0 & \text{otherwise.} \end{cases}$$

Note that by definition of NV (resp. EX, BC), we have $NV \subset NV_\infty$ (resp. $EX \subset EX_\infty$, $BC \subset BC_\infty$).

4.2 Relationships among the Inference Classes

The results of this section determine the lattice of inclusions among the inference classes NV, EX, BC, NV_∞, EX_∞, PEX_∞, BC_∞, and PBC_∞. As was previously mentioned, the standard inference notions NV, EX, and BC can easily be translated into inference notions about the recursive rational functions. We shall avoid this formalism, taking for example the assertion "$NV \subset NV_\infty$" to mean that the *analogue* of NV (in the framework of functions on $I_\mathbb{Q}$) is contained in NV_∞.

It is easy to see that restricting approximate-inference machines to output only indices of total functions results in strictly less powerful inference notions:

THEOREM 4.2.1. $PEX_\infty \subset EX_\infty$ and $PEX_\infty \subset PBC_\infty$.

PROOF. $PEX_\infty \subset EX_\infty$ is trivial. To see that $PEX_\infty \subset PBC_\infty$, let $C \in PEX_\infty$ via G, and define a PBC_∞-inference machine $\hat{G}$ as follows: at stage n, $\hat{G}$ outputs an index for $\Phi_{G((f(q_0),f(q_1),\dots,f(q_{n-1})))}$ patched with the values $\{(q_0, f(q_0)), \dots, (q_{n-1}, f(q_{n-1}))\}$. Then

$$\lim_{n\to\infty} \left\| f - \Phi_{\hat{G}((f(q_0),f(q_1),\dots,f(q_{n-1})))} \right\|_\infty = 0,$$

so that f is PBC_∞-inferred. □

This proof yields the following:

COROLLARY 4.2.2. $EX_\infty \subset BC_\infty$.

THEOREM 4.2.3. $PBC_\infty \subset BC_\infty$ and $PBC_\infty \subset NV_\infty$.

PROOF. $PBC_\infty \subset BC_\infty$ is trivial. To see that $PBC_\infty \subset NV_\infty$, let $C \in PBC_\infty$ via G, and let $f \in C$. Then

$$\lim_{n\to\infty} |f(q_n) - \Phi_{G(\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle)}(q_n)| = 0,$$

so that f is NV_∞-inferred. □

Interpolation of known function values is a natural inference procedure in the setting of recursive rational functions, and so we have

THEOREM 4.2.4. $RUC \in PBC_\infty$.

PROOF. Let $f \in RUC$. At stage n, having received $\langle f(q_0), f(q_1), \ldots, f(q_{n-1})\rangle$, simply output (an index of) the linear interpolation f_n of these n values. The functions $\{f_n\}$ are uniformly continuous on $I_\mathbb{Q}$, and converge pointwise to f, therefore $f_n \to f$ uniformly. □

By the above three theorems, we have

COROLLARY 4.2.5. $RUC \in BC_\infty$, $RUC \in NV_\infty$.

However, the interpolation method cannot be used to EX_∞-infer RUC, since it requires changing indices infinitely often. In fact, there is *no* algorithm which will EX_∞-infer RUC. The idea of the proof is essentially the same as that in the proof that $REC \notin EX$.

THEOREM 4.2.6. $RUC \notin EX_\infty$.

PROOF. Suppose G EX_∞-infers RUC. We construct an $f \in RUC$ on which G changes its mind infinitely often, for a contradiction. Let $\{a_i\}$ be an increasing, computable sequence of rationals with limit 1, and set $f(0) = f(1) = 0$. We use G to define f on all of $I_\mathbb{Q}$ as follows. At stage n, if $q_n < a_n$, do nothing. Otherwise, $q_n \in [a_k, a_{k+1})$ for some $k \geq n$. Extend f in two ways: f_1 will be f extended by 0, and f_2 is f extended

with a "hat" of height $1/k$ in $[a_k, a_{k+1})$. Both f_1 and f_2 are in *RUC*, and differ by more than $\frac{1}{2k}$ on a positive interval. Thus, there is a stage greater than n at which G outputs different indices on f_1 and f_2. Extend f by whichever of f_1, f_2 forces G to output a different index than G output on f at stage $n-1$. □

As a corollary to Theorem 4.2.2 and Theorem 4.2.6, we have

COROLLARY 4.2.7. $EX_\infty \subsetneq BC_\infty$.

In addition, although $PEX = PBC$, we have shown

COROLLARY 4.2.8. $PEX_\infty \subsetneq PBC_\infty$.

The interpolation procedure used above also does not suffice to infer an arbitrary *RC* function. We show that *RC* is not BC_∞- or NV_∞-inferable. We will need a way for an inference machine G to use its own guesses as inputs:

DEFINITION 4.2.9. For any G, $\sigma = \langle a_0, \ldots, a_{m-1}\rangle$, we define $T_{G,\sigma}$ on $I_{\mathbb{Q}}$ by recursion as follows

$$T_{G,\sigma}(q_n) = \begin{cases} G(\langle a_0, \ldots, a_{n-1}\rangle), & \text{if } n < m, \\ G(\langle a_0, \ldots, a_{m-1}, Rnd(T_{G,\sigma}(q_m)), \ldots, \\ \quad Rnd(T_{G,\sigma}(q_{n-1}))\rangle), & \text{otherwise,} \end{cases}$$

where

$$Rnd(x) = \begin{cases} 0, & \text{if } x < 1/2, \\ 1, & \text{otherwise.} \end{cases}$$

THEOREM 4.2.10. $RC \notin NV_\infty$.

PROOF. By way of contradiction, we suppose that $RC \in NV_\infty$ via G and show that we can construct a $\hat{G}$ to NV-infer $\mathbb{Q}SET$. Let $0 < r < 1$ be any computable irrational, and let $\{r_i\}$ be any increasing computable sequence of computable irrationals in $[0,1]$ with limit r. Note that if $A \in \mathbb{Q}SET$, then $\hat{A} \in RC$, where we define $\hat{A}$ by

$$\hat{A}(q) = \begin{cases} 1, & \text{if } q \in (r_i, r_{i+1}) \text{ and } A(q_i) > 1/2, \\ 0, & \text{otherwise.} \end{cases}$$

Construct $\hat{G}$ as follows. On input $\langle f(q_0), \ldots, f(q_{n-1})\rangle$ (from $f \in \mathbb{Q}REC$), output $T_{G,\langle \hat{f}(q_0), \ldots, \hat{f}(q_{m-1})\rangle}(q_n)$ for the largest m such that we can compute $\hat{f}|_m$ from $f|_n$. Then

if $A \in \mathbb{Q}SET$, then $\hat{A} \in RC$, so by hypothesis, G outputs guesses within 1/3, say, of $\hat{A}(q_n)$ for n greater than some N. Let M denote the least natural number greater than N such that $\hat{A}|_N$ can be computed from $A|_M$. Then if $A \in \mathbb{Q}SET$, for stages $n > M$, $\hat{G}$ outputs $A(q_n)$. □

Note that A can also be computed, uniformly, from $\hat{A}$, by taking $A(q_i) = \hat{A}(q_k)$, where k is the least such that $q_k \in (r_i, r_{i+1})$. Thus, there is a primitive recursive function ϕ such that if e is an index for $\hat{A}$, then $\phi(e)$ is an index for A. We will make use of ϕ in the following theorem.

THEOREM 4.2.11. $RC \notin BC_\infty$.

PROOF. As above, we suppose that $RC \in BC_\infty$ via G, and show that we can construct a $\hat{G}$ to BC-infer $\mathbb{Q}SET$. Let r, $\{r_i\}$ be as above, and for f, define $\hat{f}$ as previously. Now let $f \in \mathbb{Q}REC$. We construct $\hat{G}$ as follows. On input $\langle f(q_0), \ldots, f(q_{n-1})\rangle$, output $\phi G(\langle \hat{f}(q_0), \ldots, \hat{f}(q_{m-1})\rangle)$, for the largest m such that we can compute $\hat{f}|_m$ from $f|_n$. Thus, if $A \in \mathbb{Q}SET$, then $\hat{A} \in RC$, so by hypothesis, G outputs indices of functions that are everywhere within 1/3, say, of $\hat{A}$ for n greater than some N. Let M denote the least natural number greater than N such that $\hat{A}|_N$ can be computed from $A|_M$. But then, $\hat{G}$ outputs indices of A for $n > M$. □

The notions BC_∞, EX_∞, and NV_∞ are the $I_{\mathbb{Q}}$-domain generalizations of BC, EX, and NV, as we show in the next few theorems. Note that $NV = PEX$ and that $PEX \subset PEX_\infty$ by definition, whence $NV \subset PEX_\infty$. For completeness, however, we give the following direct proof:

THEOREM 4.2.12. $NV \subset PEX_\infty$.

PROOF. Suppose $\mathcal{C} \in NV$ via G, and $f \in \mathcal{C}$. We define a PEX_∞-inference machine H for $\mathcal{C}$ as follows. At stage n, output an index for the function h such that $h(q_i) = f(q_i)$ for each $i < n$ and $h(q_m) = G(\langle h(q_0), h(q_1), \ldots, h(q_{m-1})\rangle)$ for each $m \geq n$. Since G

makes only finitely many incorrect guesses, the sequence of indices output is eventually constant (and is in fact an index for f). $\square$

THEOREM 4.2.13. $BC \subsetneq BC_\infty$.

PROOF. Containment is clear. To show that it is strict, we show that $RUC \notin BC$. We suppose that $RUC \in BC$, and then show that $\mathbb{Q}REC \in BC$, a contradiction. Let $RUC \in BC$ via G; we will define a recursive machine H which BC-infers every recursive function on rationals. Let f be a fixed recursive function. First recall that for any rational function f, we may define an associated continuous rational function $\hat{f} \in RUC$ by defining $\hat{f}(0) = 0$ and $\hat{f}(1/i) = f(q_{i-1})/i$ for $i > 0$ and defining $\hat{f}$ for any other rational point by linear interpolation. Then the values of $\hat{f}(q_i)$ may be computed for $q_i \geq 1/n$ from $f(q_0), f(q_1), \dots, f(q_{n-1})$. Let $s(n)$ be the least s such that $q_{s+1} < \frac{1}{n}$, so that $\lim_{n\to\infty} s(n) = \infty$. Now let $E(\langle f(q_0), f(q_1), \dots, f(q_{n-1})\rangle) = G(\langle \hat{f}(q_0), \hat{f}(q_1), \dots, \hat{f}(q_{s(n)})\rangle)$. Since $\hat{f} \in RUC$, we have by assumption that $e_n = E(\langle f(q_0), f(q_1), \dots, f(q_{n-1})\rangle)$ is an index of $\hat{f}$ for all sufficiently large n. It is now straightforward to compute from e_n an index $H(\langle f(q_0), f(q_1), \dots, f(q_{n-1})\rangle)$ for the function which recovers f from $\hat{f}$. That is, if e_n is an index for the function $\hat{f}_n$, then $H(\langle f(q_0), f(q_1), \dots, f(q_{n-1})\rangle)$ is an index for the function f_n such that $f_n(q_{i-1}) = i\hat{f}_n(1/i)$. Thus H infers the arbitrary recursive function f. $\square$

In the previous literature (see Odifreddi [17], Angluin and Smith [1], and Case and Smith [6]), it has been shown that $NV \subsetneq EX \subsetneq BC$. The next few theorems show that the inductive inference hierarchy is no longer linear when we consider the approximate inference classes, since NV_∞ and BC_∞ are incomparable.

THEOREM 4.2.14. $EX \not\subseteq NV_\infty$.

PROOF. Blum and Blum [4] show that the class $\{f : \phi_{f(0)} = f\}$ is EX-inferable. It is easy to see that $C = \{f : 1/f(0) \in \mathbb{N} \text{ and } \Phi_{1/f(0)} = f\}$ is a member of (the analogue of) EX, in the context of recursive rational functions. We show that $C \notin NV_\infty$. For a

contradiction, suppose that $\mathcal{C} \in NV_\infty$ via G. By the S^m_n Theorem, we may define $\Phi_{t(e)}$ as follows. Let $\Phi_{t(e)}(0) = 1/e$. If $G(\langle\Phi_{t(e)}(0)\rangle) < 1/2$, let $\Phi_{t(e)}(1) = 1$, otherwise let $\Phi_{t(e)}(1) = 0$. Continue this process by recursion, so that, for each n, $\Phi_{t(e)}(n+1) = 1$, if $G(\langle\Phi_{t(e)}(q_0), \Phi_{t(e)}(q_1), \dots, \Phi_{t(e)}(q_n)\rangle) < 1/2$, and $\Phi_{t(e)}(n+1) = 0$, otherwise. Observe that $\Phi_{t(e)}$ is a total recursive function for all e. By the Fixed-Point Theorem, there is an index e s.t. $\Phi_{t(e)} = \Phi_e$. Then $\Phi_e \in \mathcal{C}$, but

$$|\Phi_e(q_n) - G(\langle\Phi_e(q_0), \Phi_e(q_1), \dots, \Phi_e(q_{n-1})\rangle)| \geq 1/2$$

for all n, a contradiction. □

Corollary 4.2.15. *$EX \not\subset PBC_\infty$, $\mathbb{Q}REC \notin NV_\infty$, $PBC_\infty \subsetneq BC_\infty$, $EX_\infty \not\subset NV_\infty$, and $BC_\infty \not\subset NV_\infty$.*

Proof. $EX \not\subset PBC_\infty$, since $PBC_\infty \subset NV_\infty$. $\mathbb{Q}REC \notin NV_\infty$, since otherwise $EX \subset NV_\infty$. $PBC_\infty \subsetneq BC_\infty$, since if $PBC_\infty = BC_\infty$, then $BC_\infty \subset NV_\infty$ by Theorem 4.2.3, so that $EX \subset NV_\infty$, contradicting Theorem 4.2.14. Finally, since $EX \subset EX_\infty$ and $EX \subset BC \subset BC_\infty$, the last two assertions are true. □

Proposition 4.2.16. *$BC \not\subset EX_\infty$.*

Proof. The proof of $EX^1 - EX \neq \emptyset$ in Blum and Blum [4] also shows that $EX^1 - EX_\infty \neq \emptyset$. □

The next theorem, along with the previous corollary, shows that NV_∞ and BC_∞ are incomparable.

Theorem 4.2.17. *$NV_\infty \not\subset BC_\infty$.*

Proof. Let $\mathcal{C} = \{f \in \mathbb{Q}REC : \lim_{n\to\infty} f(q_{2n}) = 1/2, f(q_{2n}) < 1/2 \Rightarrow f(q_{2n+1}) = 0, \text{ and } f(q_{2n}) \geq 1/2 \Rightarrow f(q_{2n+1}) = 1\}$. $\mathcal{C} \in NV_\infty$ by the following algorithm: at stages $2n$, simply output $1/2$, and at stages $2n+1$, use the $f(q_{2n})$ to predict whether the next value will be 0 or 1. We show that if $\mathcal{C} \in BC_\infty$ (via G, say), then $\mathbb{Q}SET \in BC$, a contradiction. We define a machine H which BC-infers $\mathbb{Q}SET$ as follows. For a

recursive set of rationals A, define $f_A(q_{2k}) = 1/2 + 1/k$ if $q_k \in A$, $f_A(q_{2k}) = 1/2 - 1/k$ if $q_k \notin A$, and $f_A(q_{2k+1}) = \chi_A(k)$. It is clear that $f_A \in \mathcal{C}$ for any A, so that by hypothesis f_A will be BC_∞-infered by G.

Now, if $A \in \mathbb{Q}REC$, on input $\langle \chi_A(q_0), \chi_A(q_1), \dots, \chi_A(q_{n-1})\rangle$, we may compute $\langle f_A(q_0), f_A(q_1), \dots, f_A(q_{2n-1})\rangle$ as defined above and then compute

$$e_n = G(\langle f_A(q_0), f_A(q_1), \dots, f_A(q_{2n-1})\rangle).$$

Since $f_A \in \mathcal{C}$, there is an N such that for all $n > N$, we have $\|f_A - \Phi_{e_n}\| < 1/2$. Thus for $n > N$ and any q, $f_A(q) = 1$ if and only if $\Phi_{e_n}(q) > 1/2$. Now use the S_n^m Theorem to compute from the program e_n a program $H(\langle \chi_A(q_0), \chi_A(q_1), \dots, \chi_A(q_{n-1})\rangle) = a_n$ so that

$$\Phi_{a_n}(q_k) \quad = \quad \begin{cases} 1, & \text{if } \Phi_{e_n}(q_{2k+1}) > 1/2, \\ 0, & \text{otherwise.} \end{cases}$$

Since $\chi_A(q_k) = f_A(q_{2k+1})$, it follows that H BC-identifies the arbitrary recursive set A. $\square$

COROLLARY 4.2.18. *$PBC_\infty \subsetneq NV_\infty$, $\mathbb{Q}REC \notin BC_\infty$.*

We have shown that the relationships among the approximate inference notions are not always analogous to the relationships among the notions of standard inference. In the sequel, we explore the reasons for this.

4.3 The Extended Inference Hierarchy

Figure 4.1 illustrates the inclusions derived in the previous section among the various inference notions. When we add the notions NV_∞ and BC_∞ to the picture, the inference hierarchy is no longer linear, and as mentioned earlier, analogues of some theorems of "standard" inference no longer hold. We offer a heuristic argument why this is to be expected.

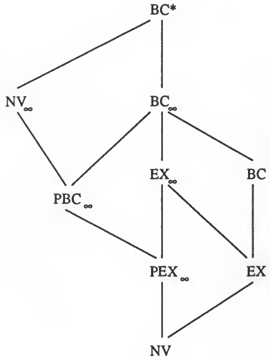

Figure 4.1: The extended inference hierarchy.

In the "standard" inference setting, it is easily shown that $NV \subsetneq BC$, but that $NV = PBC$. The proof that $PBC \subset NV$ carries over to approximate inference. The reverse inclusion does not. Consider the usual proof that $NV \subset PBC$. Given $\mathcal{C} \in NV$ (via M, say), and $f \in \mathcal{C}$, at stage n we output a program M_n which on input $k < n$ computes $M(f(q_0), f(q_1), \ldots, f(q_k))$, on input $k = n$ computes $M(\langle f(0), f(1), \ldots, f(n-1), M(\langle f(0), \ldots, f(n-1)\rangle)\rangle)$, and so on by recursion. Since M is completely accurate in its guesses from some stage N onward, the program M_n computes f if $n > N$, and so $\mathcal{C} \in PBC$.

Suppose we try to translate this proof to the new inference setting, as follows: given $\mathcal{C} \in NV_\infty$ via G, and $f \in \mathcal{C}$, use G_n (defined analogously to M_n) as a guess for f at stage n. Unfortunately, this does not work. Roughly speaking, if at stage n, G's guess for $f(q_n)$ differs from $f(q_n)$ by $\varepsilon > 0$, however small, then we expect that this error will not only propagate, but *worsen* as we ask G_n to compute on

inputs q_k for large k, so that for $k \gg n$, $|G_n(q_k) - f(q_k)| \gg 0$. In particular, we set $G_n(q_{n+1}) = G(\langle f(q_0), \ldots, f(q_{n-1}), G_n(q_n)\rangle)$, but since $G_n(q_n) \neq f(q_n)$, we don't know that $G_n(q_{n+1}) - f(q_{n+1}) < \varepsilon$. In fact, even assuming a strong continuity property for G, such as

$$\|\vec{x} - \vec{y}\| < \varepsilon \Rightarrow \|G(\vec{x}) - G(\vec{y})\| < \varepsilon$$

we may only conclude that $|G_n(q_{n+1}) - f(q_{n+1})| < 2\varepsilon$.

Of course, in the proof that $NV \subset PBC$, we might expect M_n's error to propagate in the same way. The difference is that M is accurate from some *finite* stage onward, so that from this stage on, there is no error to propagate. In contrast, we cannot be sure that G's guesses are *ever* completely accurate, only that they get "better" as time goes on. Thus, every G_n magnifies the error $|G(f(q_0), f(q_1), \ldots, f(q_{n-1})) - f(n)|$ when input q_k, $k \gg n$, so that we cannot expect that f is the uniform limit of the sequence $\{G_n\}$.

Recall that all of the standard inference notions have the following "finitistic" component which is lacking in the notions of approximate inference given thus far. Using NV as an example, if f is inferred via M, then by some (finite) stage N, M will ever after predict correctly the next value to be input from f (although M does not "know" when this stage N occurs). Thus, unlike the approximate inference methods defined up to now, there is a criterion for accuracy which is met at some finite stage. The next chapter gives methods for approximate inference which also have this feature.

CHAPTER 5
WEAKER NOTIONS OF APPROXIMATE INFERENCE

5.1 The "Epsilon" Inference Criteria

Recall the motivation for approximate inference from the previous chapter. $f : [0,1] \longrightarrow [0,1]$ is continuous, and we are allowed to ask for a sequence of values $f(x)$, say $f(0)$, $f(1)$, $f(1/2)$, $f(1/3)$, $f(2/3)$, and so on. As before, after we ask for the n^{th} value, we may construct an approximation f_n to the graph of f, by some form of continuous interpolation. At some point in this procedure, we may be satisfied that we are "close enough" for our particular purposes, and so, no longer wish to continue to build the approximations $\{f_n\}$, settling instead on, say, $g = f_N$ for some fixed N in all later computations. If we have in mind that we wish to be within, say, $\varepsilon = 0.00001$ of $f(x)$ for all x, we can keep asking for new values of $f(x)$, checking these to make sure that $|g(x) - f(x)| < \varepsilon$. If we eventually happen upon an x for which this does not hold, we may then update our interpolant g. However, since the f_n's converge uniformly to f, we will only have to make finitely many such updates. Thus, analogously to the standard inference notions, this procedure allows us to meet our inference criterion by some finite stage, although we cannot in general determine when this stage occurs.

We now modify the notions of approximate inference accordingly to formalize this idea. The BC_∞ inference criterion is essentially one of uniform convergence. Thus, a natural weakening of this criterion is to require convergence of the guesses only to within ε for a fixed $\varepsilon > 0$. We may similarly weaken other approximate inference criteria. In the sequel, we take ε to be rational, for purposes of computability.

DEFINITION 5.1.1. Let $\varepsilon > 0$. A class $\mathcal{C}$ of recursive rational functions is *next-value approximable to within* ε ($\mathcal{C} \in NV_\varepsilon$) if there is a $G : [I_{\mathbb{Q}}]^{<\omega} \longrightarrow I_{\mathbb{Q}}$ such that for every $f \in \mathcal{C}$

$$\limsup_n |f(q_n) - G(\langle f(q_0), \ldots, f(q_{n-1})\rangle)| < \varepsilon.$$

DEFINITION 5.1.2. Let $\varepsilon > 0$. A class $\mathcal{C}$ of recursive rational functions is *explanatorily approximable to within* ε ($\mathcal{C} \in EX_\varepsilon$) if there is a $G : [I_{\mathbb{Q}}]^{<\omega} \longrightarrow \mathbb{N}$ such that for every $f \in \mathcal{C}$, there is an index e such that

- Φ_e is total,
- for all but finitely many n, $G(\langle f(q_0), \ldots, f(q_{n-1})\rangle) = e$, and

$$\limsup_n |f(q_n) - \Phi_e(q_n)| < \varepsilon.$$

The definition for $\mathcal{C} \in PEX_\varepsilon$ is similar, but G must output only indices for *total* functions.

DEFINITION 5.1.3. A class $\mathcal{C}$ of recursive rational functions is *behaviorally correctly approximable to within* ε ($\mathcal{C} \in BC_\varepsilon$) if there is a recursive $G : [I_{\mathbb{Q}}]^{<\omega} \longrightarrow \mathbb{N}$ such that for every $f \in \mathcal{C}$,

$$\limsup_n \|f - \Phi_{G(\langle f(q_0), \ldots, f(q_{n-1})\rangle)}\|_\infty < \varepsilon.$$

So, G BC_ε-infers $\mathcal{C}$ if for each $f \in \mathcal{C}$, $\exists N\ \forall n > N\ G(\langle f(q_0), f(q_1), \ldots, f(q_n)\rangle$ is an index of a recursive function whose values are strictly within ε if f's values. The definition for $\mathcal{C} \in PBC_\varepsilon$ is similar, but G must output only indices for *total* functions.

Each of the above criteria yields a hierarchy of inference classes parameterized by $\varepsilon > 0$. Clearly, for $\varepsilon > 1/2$, the class $\mathcal{I}_\varepsilon$ contains $\mathbb{Q}REC$ for every inference notion $\mathcal{I} \in \{NV, EX, PEX, BC, PBC\}$, so the hierarchies collapse above $\varepsilon = 1/2$. We will show that the hierarchies do not collapse below $\varepsilon = 1/2$, and are in fact strictly monotone. In the sequel, δ, ε range over rational numbers in $[0,1]$. Let δf be the function obtained from f by pointwise multiplication by δ, and for a class of functions

$\mathcal{C}$, let $\varepsilon\mathcal{C}$ denote the class $\{\delta f | f \in \mathcal{C}, 0 \leq \delta \leq \varepsilon\}$. It is easy to see that for any $\varepsilon > 0$, if $0 < \delta < 2\varepsilon$, then $\delta\mathbb{Q}REC$ is an element of each $\mathcal{I}_\varepsilon$ for $\mathcal{I} \in \{NV, EX, PEX, BC, PBC\}$.

We begin by showing that these new inference notions are strictly weaker than the notions of approximate inference, in the sense that the new notions infer more classes of functions.

THEOREM 5.1.4. For any $\varepsilon > 0$, $NV_\infty \subsetneq NV_\varepsilon$.

PROOF. Note that $\varepsilon\mathbb{Q}REC \in NV_\varepsilon$. We show that $\varepsilon\mathbb{Q}REC \notin NV_\infty$: for any G (which potentially NV_∞-infers $\varepsilon\mathbb{Q}REC$), we find an f in $\varepsilon\mathbb{Q}REC$ not inferred by G. Let Ψ be defined by

$$\Psi(q_n) = \begin{cases} \varepsilon, & \text{if } G(\langle \Psi(q_0), \Psi(q_1), \ldots, \Psi(q_{n-1}) \rangle) < \frac{\varepsilon}{2}, \\ 0, & \text{otherwise.} \end{cases}$$

Since Ψ is recursive, and $\|\Psi\|_\infty \leq \varepsilon$, Ψ is an element of $\varepsilon\mathbb{Q}REC$. But

$$\limsup_n |\Psi(q_n) - G(\langle \Psi(q_0), \Psi(q_1), \ldots, \Psi(q_{n-1}) \rangle)| \geq \frac{\varepsilon}{2},$$

so $\varepsilon\mathbb{Q}REC$ is not NV_∞-inferred by G. □

THEOREM 5.1.5. For $\varepsilon > 0$, $EX_\infty \subsetneq EX_\varepsilon$.

PROOF. Note that $\varepsilon\mathbb{Q}REC \in EX_\varepsilon$. We show that $\varepsilon\mathbb{Q}REC \notin EX_\infty$: Suppose that $\varepsilon\mathbb{Q}REC \in EX_\infty$. Then $\varepsilon\mathbb{Q}SET \in EX_\infty$ (via G, say), since $\mathbb{Q}SET \subset \mathbb{Q}REC$. But then if $f \in \mathbb{Q}SET$, feed εf to G. Let e_n be the index output by G at stage n, i.e. $e_n = G(\langle \varepsilon f(q_0), \ldots, \varepsilon f(q_{n-1}) \rangle)$. Then, since G EX_∞-infers εf, there is an index e and a stage N so that for each $n > N$, $e_n = e$, and

$$\lim_k |\Phi_e(q_k) - \varepsilon f(q_k)| < \frac{\varepsilon}{2}.$$

We use this to EX_∞-infer $\mathbb{Q}SET$ as follows. For each n we output an index of the function Ψ_n defined by

$$\Psi_n(x) = \begin{cases} 1, & \text{if } \Phi_{e_n}(x) \downarrow > \frac{\varepsilon}{2}, \\ 0, & \text{if } \Phi_{e_n}(x) \downarrow \leq \frac{\varepsilon}{2}, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Since $e_n = e$ after stage N, with $|\Phi_e(q_k) - \varepsilon f(q_k)| < \frac{\varepsilon}{2}$ for sufficiently large k, then because $f \in \mathbb{Q}SET$, we have $|\Psi_n(q_k) - f(q_k)| = 0$ for such k. But this procedure is uniform in f, so it defines a EX_∞-inference machine for $\mathbb{Q}SET$, a contradiction. Hence $\varepsilon\mathbb{Q}REC \in EX_\varepsilon - EX_\infty$. □

THEOREM 5.1.6. For $\varepsilon > 0$, $BC_\infty \subsetneq BC_\varepsilon$.

PROOF. Note that $\varepsilon\mathbb{Q}REC \in BC_\varepsilon$. We show that $\varepsilon\mathbb{Q}REC \notin BC_\infty$: Suppose that $\varepsilon\mathbb{Q}REC \in BC_\infty$. Then $\varepsilon\mathbb{Q}SET \in BC_\infty$ (via G, say), since $\mathbb{Q}SET \subset \mathbb{Q}REC$. But then if $f \in \mathbb{Q}SET$, feed εf to G. Let e_n be the index output by G at stage n, i.e. $e_n = G(\langle \varepsilon f(q_0), \dots, \varepsilon f(q_{n-1})\rangle)$, then since G BC_∞-infers εf, we have

$$\lim_n \|\Phi_{e_n} - \varepsilon f\|_\infty = 0,$$

so there is a stage N so that for each $n > N$,

$$\|\Phi_{e_n} - \varepsilon f\|_\infty < \frac{\varepsilon}{2}.$$

We use this to BC_∞-infer $\mathbb{Q}SET$ as follows. For each n we output an index of the function Ψ_n defined by

$$\Psi_n(x) = \begin{cases} 1, & \text{if } \Phi_{e_n}(x)\downarrow > \frac{\varepsilon}{2}, \\ 0, & \text{if } \Phi_{e_n}(x)\downarrow \le \frac{\varepsilon}{2}, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Since the $\|\Phi_{e_n} - \varepsilon f\|_\infty < \frac{\varepsilon}{2}$ after stage N, if $f \in \mathbb{Q}SET$, then $\Psi_n = f$ after stage N. But this procedure is uniform in f, so it defines a BC_∞-inference machine for $\mathbb{Q}SET$, a contradiction. Hence $\varepsilon\mathbb{Q}REC \in BC_\varepsilon - BC_\infty$. □

The preceding two proofs, mutadis mutandis, yield analogous results for PEX_∞ and PBC_∞:

COROLLARY 5.1.7. For $\varepsilon > 0$, $PEX_\infty \subsetneq PEX_\varepsilon$.

COROLLARY 5.1.8. For $\varepsilon > 0$, $PBC_\infty \subsetneq PBC_\varepsilon$.

Above, we exhibited, for each ε, a class $\mathcal{C}_\varepsilon$ such that $\mathcal{C}_\varepsilon \in EX_\varepsilon - EX_\infty$. We can further show that there is a class $\mathcal{C}$ with $\mathcal{C} \in \cap_\varepsilon EX_\varepsilon - EX_\infty$, and that $\mathcal{C} \in EX_\varepsilon$ *uniformly* in ε. This class $\mathcal{C}$ is just RUC.

THEOREM 5.1.9. $RUC \in EX_\varepsilon$ for all ε.

PROOF. The machine G_ε which accomplishes this is a simple variant of the linear interpolation procedure used previously. Fix $f \in \mathbb{Q}REC$, let L be, initially, the zero function, and let L_n denote the linear interpolation of $\langle f(q_0), \ldots, f(q_{n-1})\rangle$. We construct G_ε as follows: at stage n, if $L(q_{n-1})$ is not within ε of $f(q_{n-1})$, set $L = L_n$. Output L.

Now, if $f \in RUC$, eventually, we will reach a stage after which all of the linear interpolants L_n are everywhere within ε of f. G_ε will stabilize on the first such interpolant. □

In the case of NV_∞ it is straightforward to show that there is a $\mathcal{C}$ such that $\mathcal{C} \in \cap_\varepsilon NV_\varepsilon - NV_\infty$, although not necessarily uniformly. For each $n > 0$ and x, let $s_n(x) = \frac{1}{2^{n+1}}x + \frac{1}{2^n}$. Then for fixed n, for any $f \in \mathbb{Q}REC$, and all $x \in I_{\mathbb{Q}}$, we may define $S_{f,n}(x) = s_n(f(x))$. Now, enumerate $\mathbb{Q}SET$ arbitrarily as $\{f_1, f_2, \ldots\}$. In the next two theorems, let $\mathcal{C} = \{S_{f_i,i}\}$, and let $\nu(n)$ be an effective enumeration of $\mathbb{N}$, in which each natural number appears infinitely often.

THEOREM 5.1.10. $NV_\infty \subsetneq \cap_\varepsilon NV_\varepsilon$.

PROOF. We show that $\mathcal{C} \in NV_{1/k}$ for any $k > 1$, but that $\mathcal{C} \notin NV_\infty$. Fix $k > 1$, and let $\hat{f}_i = S_{f_i,i}$. Then $\mathcal{C} \in NV_{1/k}$ via G_k, defined as follows (input $f \in \mathbb{Q}REC$):

Stage n:

Step 1: Find $l > 0$ such that $f(q_{n-1}) \in [\frac{1}{2^l}, \frac{1}{2^{l-1}})$ (if no such l exists— i.e $f(q_n)$ is 0 or 1— output 1, and goto stage $n+1$).

Step 2: If $l < k$, output $\hat{f}_l(q_n)$, otherwise, output $\frac{1}{2^l}$.

End of Construction.

Clearly, if $f \in \mathcal{C}$, then G_k $NV_{1/k}$-infers f. Note that the construction of G_k is not uniform in k, since no enumeration of $\mathbb{Q}REC$ is effective. Intuitively, this is why $\mathcal{C}$ is not NV_∞-inferable. We suppose $\mathcal{C}$ *is* NV_∞-inferable, via G, say and show that we can then NV-infer $\mathbb{Q}SET$, for a contradiction. Construct M as follows (input $f \in \mathbb{Q}REC$, and let $e = 0$):

Stage n:

Step 1: Compute $v_n = G(S_{f,\nu(e)}|_n)$.

Step 2: If v_n is closer to $\frac{1}{2^{\nu(e)}}$ than to $\frac{1}{2^{\nu(e)+1}}$ output 1, else output 0.

Step 3: If

$$|G(S_{f,\nu(e)}|_{n-1}) - S_{f,\nu(e)}(q_{n-1})| \geq \frac{1}{2^{\nu(e)+2}},$$

increment e.

End of Construction.

Now if $A \in \mathbb{Q}SET$, then by the construction, since G infers $\mathcal{C}$, M infers A. Given $f = f_i \in \mathbb{Q}SET$, let $\hat{f} = f_i = S_{f,i}$. Then $\hat{f} \in \mathcal{C}$, and $\hat{f}(n) = \frac{1}{2^i}$ or $\frac{1}{2^i} + \frac{1}{2^{i+1}}$ for each n. Let N be large enough so that

$$|G(\langle \hat{f}(q_0), \ldots, \hat{f}(q_{n-1})\rangle) - \hat{f}(q_n)| < \frac{1}{2^{i+2}}$$

for all $n > N$. There are two cases to consider.

Case 1: There is a stage $m \geq N$ where $\nu(e) = i$. Then for every $n \geq m$

$$G(S_{f,\nu(e)}|_{n-1}) = G(\langle \hat{f}(q_0), \ldots, \hat{f}(q_{n-2})\rangle),$$

and $S_{f,\nu(e)}(q_{n-1}) = \hat{f}(q_{n-1})$, so that the difference computed in Step 3 is less than $\frac{1}{2^{i+1}}$, which means that $\nu(e)$ remains equal to i thereafter. Thus, for all $n \geq m$, $|v_n - \hat{f}(q_n)| < \frac{1}{2^{i+2}}$, so that by Step 2 of the construction, since $f \in \mathbb{Q}SET$, $M(\langle f(q_0), \ldots, f(q_{n-1})\rangle) = f(q_n)$ for all $n \geq m$.

Case 2: There is no stage $m \geq N$ where $\nu(e) = i$. Since $\nu(e) = i$ for infinitely many e, this means that e remains fixed from some stage m onwards. But this implies, by Step 3, that G is successfully inferring $S_{f,\nu(e)}(q_n)$ for all $n \geq m$. It follows, as in the previous case, that M infers f. Note that in this case, G is inferring $S_{f,\nu(e)}$ even though $S_{f,\nu(e)}$ may not be in $\mathcal{C}$. □

We remark that the above proof technique is independent of the enumeration of $\mathbb{Q}REC$ chosen, and therefore does not work to show that $BC_\infty \subsetneq \cap_\varepsilon BC_\varepsilon$. To see this, enumerate $\mathbb{Q}SET$ as follows: let $f_i = \phi_i$ if i is the index of a total recursive function which is not identically zero, otherwise let $f_i = \lambda x.0$. Then the class $\mathcal{C} = \{S_{f_i,i}\}$ is *EX*-inferable (and therefore BC_∞-inferable) by the machine which at stage n outputs

- an index of $\lambda x.0$ if input the constant sequence $\langle \frac{1}{2^{e+1}}, \ldots, \frac{1}{2^{e+1}} \rangle$ of length n,
- the index e if input any other sequence $\langle a_0, \ldots, a_{n-1} \rangle$ such that $a_i \in [\frac{1}{2^e}, \frac{1}{2^e} + \frac{1}{2^{e-1}})$ for all i,
- 0, otherwise.

We do not know of a proof for $BC_\infty \subsetneq \cap BC_\varepsilon$, but we conjecture that the statement is true.

5.2 Monotonicity of the "Epsilon" Hierarchies

By adapting the proofs of the previous section, it is easy to show that the "epsilon" hierarchies do not collapse for $\varepsilon \leq 1/2$:

THEOREM 5.2.1. For any $0 < \varepsilon < 1/2$, $2\varepsilon\mathbb{Q}REC \notin NV_\varepsilon$.

PROOF. We show that for any G (which potentially NV_ε-infers $2\varepsilon\mathbb{Q}REC$), we can find an f in $2\varepsilon\mathbb{Q}REC$ not inferred by G. Let Ψ be defined by

$$\Psi(q_n) = \begin{cases} 2\varepsilon, & \text{if } G(\langle \Psi(q_0), \Psi(q_1), \ldots, \Psi(q_{n-1}) \rangle) < \varepsilon, \\ 0, & \text{otherwise.} \end{cases}$$

Since Ψ is recursive, and $\|\Psi\|_\infty \leq 2\varepsilon$, Ψ is an element of $2\varepsilon\mathbb{Q}REC$. But

$$\limsup_n \Psi(q_n) - G(\langle\Psi(q_0), \Psi(q_1), \dots, \Psi(q_{n-1})\rangle) \geq \varepsilon,$$

so $2\varepsilon\mathbb{Q}REC$ is not NV_ε-inferred by G. □

Thus, the $\{NV_\varepsilon\}_{0<\varepsilon\leq 1/2}$ hierarchy does not collapse:

COROLLARY 5.2.2. If $0 < \varepsilon_1 < \varepsilon_2 \leq 1/2$, then $NV_{\varepsilon_1} \subsetneq NV_{\varepsilon_2}$.

PROOF. Any NV_{ε_1}-inference machine is by definition an NV_{ε_2}-inference machine, so $NV_{\varepsilon_1} \subseteq NV_{\varepsilon_2}$. This containment is strict, however, by the preceding theorem, since $2\varepsilon_1\mathbb{Q}REC \in NV_{\varepsilon_2}$, via G which simply outputs ε_1 on any input. □

Also, since NV_ε contains both PEX_ε and PBC_ε, we have:

COROLLARY 5.2.3. For $\varepsilon \leq 1/2$, $2\varepsilon\mathbb{Q}REC \notin PBC_\varepsilon$ and $2\varepsilon\mathbb{Q}REC \notin PEX_\varepsilon$.

COROLLARY 5.2.4. If $0 < \varepsilon_1 < \varepsilon_2 \leq 1/2$, then $PBC_{\varepsilon_1} \subsetneq PBC_{\varepsilon_2}$, and $PEX_{\varepsilon_1} \subsetneq PEX_{\varepsilon_2}$.

THEOREM 5.2.5. For $0 < \varepsilon \leq 1/2$, $2\varepsilon\mathbb{Q}REC \notin EX_\varepsilon$.

PROOF. We suppose $2\varepsilon\mathbb{Q}REC \in EX_\varepsilon$, for a contradiction. Then $2\varepsilon\mathbb{Q}SET \in EX_\varepsilon$ (via G, say), since $\mathbb{Q}SET \subset \mathbb{Q}REC$. But then if $f \in \mathbb{Q}SET$, feed $2\varepsilon f$ to G. Let e_n be the index output by G at stage n, i.e. $e_n = G(\langle 2\varepsilon f(q_0), \dots, 2\varepsilon f(q_{n-1})\rangle)$. Then, since G EX_ε-infers $2\varepsilon f$, there is an index e and a stage N so that for each $n > N$, $e_n = e$, and

$$\lim_k |\Phi_e(q_k) - 2\varepsilon f(q_k)| < \varepsilon.$$

We use this to EX-infer $\mathbb{Q}SET$ as follows. For each n we output an index of the function Ψ_n defined by

$$\Psi_n(x) = \begin{cases} 1, & \text{if } \Phi_{e_n}(x)\downarrow > \varepsilon, \\ 0, & \text{if } \Phi_{e_n}(x)\downarrow \leq \varepsilon, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Since $e_n = e$ after stage N, and $|\Phi_e(q_k) - 2\varepsilon f(q_k)| < \varepsilon$ for sufficiently large k, then $|\Psi_n(q_k) - f(q_k)| = 0$ for such k. But this procedure is uniform in f, so it defines a EX-inference machine for $\mathbb{Q}SET$, a contradiction. Hence $2\varepsilon\mathbb{Q}REC \notin EX_\varepsilon$. □

COROLLARY 5.2.6. If $0 < \varepsilon_1 < \varepsilon_2 \leq 1/2$, then $EX_{\varepsilon_1} \subsetneq EX_{\varepsilon_2}$.

PROOF. Any EX_{ε_1}-inference machine is by definition an EX_{ε_2}-inference machine, so $EX_{\varepsilon_1} \subseteq EX_{\varepsilon_2}$. This containment is strict, however, by the preceding corollary, since $2\varepsilon_1 \mathbb{Q}REC \in EX_{\varepsilon_2}$, via G which simply outputs an index for the constant function$f(q) = \varepsilon_1$ on any input. $\square$

In the following, if $|g - f| < \varepsilon$, we will call g an *ε-variant* of f.

THEOREM 5.2.7. For $0 < \varepsilon \leq 1/2$, $2\varepsilon\mathbb{Q}REC \notin BC_\varepsilon$.

PROOF. We suppose $2\varepsilon\mathbb{Q}REC \in BC_\varepsilon$, for a contradiction. Then $2\varepsilon\mathbb{Q}SET \in BC_\varepsilon$ (via G, say), since $\mathbb{Q}SET \subset \mathbb{Q}REC$. But then if $f \in \mathbb{Q}SET$, feed $2\varepsilon f$ to G. Let e_n be the index output by G at stage n, i.e. $e_n = G(\langle 2\varepsilon f(q_0), \ldots, 2\varepsilon f(q_{n-1})\rangle)$. Then, since G BC_ε-infers $2\varepsilon f$, we have

$$\lim_n \|\Phi_{e_n} - 2\varepsilon f\|_\infty = 0,$$

so there is a stage N so that for each $n > N$,

$$\|\Phi_{e_n} - 2\varepsilon f\|_\infty < \varepsilon.$$

Thus, for each stage $n > N$, G produces an index of an ε-variant of f. We use this fact to BC-infer $\mathbb{Q}SET$ as follows. For each n we output an index of the function Ψ_n defined by

$$\Psi_n(x) = \begin{cases} 1, & \text{if } \Phi_{e_n}(x) \downarrow > \varepsilon, \\ 0, & \text{if } \Phi_{e_n}(x) \downarrow \leq \varepsilon, \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

Since the $\|\Phi_{e_n} - 2\varepsilon f\|_\infty < \varepsilon$ after stage N, then $\Psi_n = f$ after stage N. But this procedure is uniform in f, so it defines a BC-inference machine for $\mathbb{Q}SET$, a contradiction. Hence $2\varepsilon\mathbb{Q}REC \notin BC_\varepsilon$. $\square$

COROLLARY 5.2.8. If $0 < \varepsilon_1 < \varepsilon_2 \leq 1/2$, then $BC_{\varepsilon_1} \subsetneq BC_{\varepsilon_2}$.

PROOF. Any BC_{ε_1}-inference machine is by definition an BC_{ε_2}-inference machine, so $BC_{\varepsilon_1} \subseteq BC_{\varepsilon_2}$. This containment is strict, however, by the preceding corollary, since $2\varepsilon_1 \mathbb{Q}REC \in BC_{\varepsilon_2}$, via G which simply outputs an index for the constant function $f(q) = \varepsilon_1$ on any input. □

5.3 Relationships among the Hierarchies

Note that Theorems 4.2.14 and 4.2.17 are easily modified to yield the following:

COROLLARY 5.3.1. (to 4.2.14) For any $0 < \varepsilon \leq 1/2$, $EX \not\subset NV_\varepsilon$.

COROLLARY 5.3.2. (to 4.2.17) For any $0 < \varepsilon \leq 1/2$, $NV_\infty \not\subset BC_\varepsilon$.

Most of the assertions below are just corollaries to the above, and to the theorems of the previous section. For A, B, families of subsets of REC, we use the notation $A \perp B$ to mean $A \not\subset B$ and $B \not\subset A$.

COROLLARY 5.3.3. For all $0 < \delta, \varepsilon \leq 1/2$, $NV_\varepsilon \perp BC_\delta$.

PROOF. $NV_\varepsilon \not\subset BC_\delta$: $NV_\infty \subset NV_\varepsilon$, but by 5.3.2, $NV_\infty \not\subset BC_\delta$. $BC_\delta \not\subset NV_\varepsilon$: $EX \subset BC \subset BC_\delta$, but by 5.3.1, $EX \not\subset NV_\varepsilon$. □

COROLLARY 5.3.4. For all $0 < \delta, \varepsilon \leq 1/2$, $NV_\varepsilon \perp EX_\delta$.

PROOF. $NV_\varepsilon \not\subset EX_\delta$: $NV_\infty \subset NV_\varepsilon$, but by 5.3.2, since $EX_\delta \subset BC_\delta$, $NV_\infty \not\subset EX_\delta$. $EX_\delta \not\subset NV_\varepsilon$: $EX \subset EX_\delta$, but by 5.3.1, $EX \not\subset NV_\varepsilon$. □

COROLLARY 5.3.5. If $0 < \varepsilon \leq 1/2$, $PBC_\varepsilon \subsetneq BC_\varepsilon$.

PROOF. Containment is immediate. It must be strict, since otherwise $EX \subset NV_\varepsilon$, a contradiction. □

COROLLARY 5.3.6. If $0 < \varepsilon \leq 1/2$, $PBC_\varepsilon \subsetneq NV_\varepsilon$.

PROOF. Containment is immediate. It must be strict, since otherwise $NV_\infty \subset BC_\varepsilon$, a contradiction. □

Unlike the case of PEX_∞ and PBC_∞ in the previous chapter, in the framework of weak approximate inference, we can easily modify the proof of $PEX = PBC$ to yield its analogue:

LEMMA 5.3.7. If $0 < \varepsilon \leq 1/2$, $PEX_\varepsilon = PBC_\varepsilon$.

PROOF. If $C \in PEX_\varepsilon$ via G, on input f, at stage n we patch the index $G(\langle f(q_0), \dots, f(q_n)\rangle)$ with the known values $\langle f(q_0), \dots, f(q_n)\rangle$. If $f \in C$, then this procedure serves to BC_ε-infer f. In the other direction, if $C \in PBC_\varepsilon$ via G, on input f, at stage n simply output the least index $e_i = G(\langle f(q_0), \dots, f(q_i)\rangle)$ (for $i < n$) for which $|\Phi_i(q_x) - f(q_x)| < \varepsilon$ for each $x < n$ (if no such i exists, just output e_n). □

COROLLARY 5.3.8. If $0 < \varepsilon \leq 1/2$, $PEX_\varepsilon \subsetneq EX_\varepsilon$.

PROOF. Containment is strict: otherwise, since $PEX_\varepsilon = PBC_\varepsilon$, we would have $EX \subset NV_\varepsilon$, a contradiction. □

Previously, it was noted that the proof of $EX^1 - EX \neq \emptyset$ in Blum and Blum [4] also shows that $EX^1 - EX_\infty \neq \emptyset$. In fact, it is easily modified to yield $EX^1 - EX_\varepsilon \neq \emptyset$ for $0 < \varepsilon \leq 1/2$. Thus, since $EX^1 \subsetneq BC$, the following proposition holds:

PROPOSITION 5.3.9. If $0 < \varepsilon \leq 1/2$, $EX_\varepsilon \subsetneq BC_\varepsilon$.

Figure 5.1 illustrates the inclusions derived in this section.

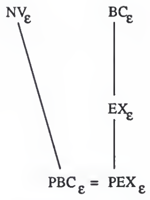

Figure 5.1: The "epsilon" hierarchy.

CHAPTER 6
INFERENCE FROM ORACLES

6.1 Inference from Generic Oracles

We now turn to approximate inference using oracles (see chapter 2 for notation and definitions). We would like to obtain analogues for approximate inference of the results in Fortnow et al. [9] characterizing oracle triviality for EX and BC, namely that

$$EX[A] = EX \Leftrightarrow \mathcal{G}(A) \text{ and } BC[A] = BC \Leftrightarrow \mathcal{G}(A),$$

where $\mathcal{G}(A)$ is the condition that either A is recursive, or $A \leq_T K$ and $A \equiv_T G$ for some generic G.

As noted previously, the crux of the argument is to show that for generic G, $EX[G*] = EX[G]$ and that $BC[G*] = BC[G]$. It appears however, that similar relationships do not hold in the cases of NV_∞ and BC_∞. In the sequel, we introduce the concept of a modulus of inference, and show that for classes $\mathcal{C}$ which can be inferred by machines M^G with recursive inference moduli, only finitely many queries are needed. This, in turn, leads to new notions of approximate inference.

We recall the following definition of genericity provided by Fortnow et al. [9]:

DEFINITION 6.1.1. $G \in 2^\omega$ is *generic* if for each Σ_1 $W \subset \{0,1\}^*$, either

- $(\exists \sigma \prec G)\ (\forall \tau \succeq \sigma)\ \tau \notin W$, or
- $(\exists \sigma \prec G)\ \sigma \in W$.

In the sequel, we will use the following formulation (see Jockusch [13] for other characterizations).

LEMMA 6.1.2. G is generic if and only if for each Π^0_1 class $\mathcal{P} \subset 2^\omega$, either

- $(\exists\sigma \prec G)\, I(\sigma) \subset \mathcal{P}$, or
- $G \notin \mathcal{P}$.

PROOF. ($\Rightarrow$): If $\mathcal{P}$ is a Π^0_1 class, then $\mathcal{P} = [T]$ for some recursive tree T (note that $\overline{T}$ is Σ_1). If $G \notin \mathcal{P}$ there is nothing to prove. Suppose that $G \in \mathcal{P}$. Then $\forall\sigma \prec G$, $\sigma \notin \overline{T}$. Thus, by genericity of G, there is a $\sigma \prec G$ so that for each $\tau \succeq \sigma, \tau \notin \overline{T}$, whence $I(\sigma) \subset \mathcal{P}$.

($\Leftarrow$): Suppose W is Σ_1, and $\forall\sigma \prec G$, $\sigma \notin W$. Now $\overline{W} \in \Pi^0_1$, so $[\overline{W}]$ is a Π^0_1 class, and $G \in [\overline{W}]$. Thus, there is a $\sigma \prec G$ with $I(\sigma) \subset [\overline{W}]$, so for all $\tau \succeq \sigma$, $\tau \notin W$. □

It was shown [9] that for generic G, $EX[G] = EX[G*]$, and $BC[G] = BC[G*]$. $EX_\infty[G] = EX_\infty[G*]$ is essentially a corollary of the first result. However, $BC_\infty[G] = BC_\infty[G*]$ does not seem to follow from any simple modification of the proof $BC[G] = BC[G*]$. We suspect that in fact, $BC_\infty[G*] \subsetneq BC_\infty[G]$, and also that $NV_\infty[G*] \subsetneq NV_\infty[G]$. However, we can show that, at least for NV_ϵ- and BC_ϵ-inference from generic oracles, no more power is obtained from an infinite number than from an arbitrary finite number of queries. The basic idea is to compute from all oracles in an interval $I(\sigma)$, where $\sigma \prec G$, as long as the various computations at any given stage are all "close" to each other. If not all the computations are close, then we ask for a longer initial segment $\sigma \prec \tau \prec G$, and start computing from oracles in $I(\tau)$. Since G is generic, we will only have to return to G finitely many times to obtain these initial segments.

DEFINITION 6.1.3. For $\tau \in 2^{<\omega}$, let S_τ denote the partial recursive function with domain $|\tau|$, defined by $S_\tau(x) = (\tau)_x$. We say that τ is $(M, \langle x_0, \ldots, x_n\rangle)$*-minimal* if $M^{S_\tau}(\langle x_0, \ldots, x_n\rangle)\downarrow$, but for all proper initial segments σ of τ, $M^{S_\sigma}(\langle x_0, \ldots, x_n\rangle)\uparrow$.

LEMMA 6.1.4. For any categorical $M^{()}$, compact $\mathcal{D} \subset 2^\omega$, and sequence $\langle a_0, \dots, a_n \rangle$, the set

$$\{M^B(\langle a_0, \dots, a_n \rangle) \mid B \in \mathcal{D}\}$$

is finite.

PROOF. For any $B \in \mathcal{D}$, since $M^B(\langle a_0, \dots, a_n \rangle) \downarrow$, there is a $\sigma_B \prec B$ so that M uses only σ_B in its computation. Thus, for all $A \in I(\sigma_B)$, $M^B(\langle a_0, \dots, a_n \rangle) = M^A(\langle a_0, \dots, a_n \rangle)$. Now, $\{I(\sigma_B) \mid B \in \mathcal{D}\}$ is an open cover of $\mathcal{D}$, so it contains a finite subcover $I(\sigma_{B_1}), \dots, I(\sigma_{B_k})$. Thus,

$$\{M^B(\langle a_0, \dots, a_n \rangle) \mid B \in \mathcal{D}\} =$$
$$\{M^{B_1}(\langle a_0, \dots, a_n \rangle), \dots, M^{B_k}(\langle a_0, \dots, a_n \rangle)\}$$

is finite. □

Note that if the relation $B \in \mathcal{D}$ is computable, then

$$\{M^B(\langle a_0, \dots, a_n \rangle) \mid B \in \mathcal{D}\}$$

is uniformly computable in $\langle a_0, \dots, a_n \rangle$.

THEOREM 6.1.5. If G is generic, then $NV_{1/k}[G*] = NV_{1/k}[G]$.

PROOF. Let $C \in NV_{1/k}[G]$ via M^G. We construct $\hat{M}$ to $NV_{1/k}[G*]$-infer C. Let $f \in \mathbb{Q}REC$, and initialize σ_0 to $\emptyset$. Then $\hat{M}$ works as follows:

Stage n:

Step 1: Output $a_n = M^{\sigma_n 0^*}(\langle f(q_0), \dots, f(q_{n-1}) \rangle)$.

Step 2: For each $\tau \succ \sigma_n$ such that τ is $(M, \langle f(q_0), \dots, f(q_{n-2}) \rangle)$-minimal compute

$$v_\tau = M^\tau(\langle f(q_0), \dots, f(q_{n-2}) \rangle).$$

Note that since M is categorical, by the above lemma, all such τ can be found effectively. Also, the list may be empty, since M might need only a proper initial segment of σ.

Step 3: If for some τ we have $|v_\tau - f(q_{n-1})| \geq 1/k$, let $\sigma_{n+1} = G|_n$, otherwise, let $\sigma_{n+1} = \sigma_n$.

End of Construction.

If $f \in \mathcal{C}$, we claim that $\hat{M}$ makes only finitely many queries to G. Since f is $NV_{1/k}$-inferred by M^G, there is an N_f so that for each $n > N_f$, $|M^G(\langle f(q_0), \ldots, f(q_{n-1})\rangle) - f(q_n)| < 1/k$. Let

$$\mathcal{P} = \{B \mid \forall n \geq N_f\ |M^B(\langle f(q_0), \ldots, f(q_{n-1})\rangle) - f(q_n)| < 1/k\}.$$

Then $\mathcal{P}$ is a Π^0_1 class, and $G \in \mathcal{P}$, so there is a $\sigma \prec G$ with $I(\sigma) \subset \mathcal{P}$. Thus, the consequent of step 3 is invoked only finitely often, for otherwise, there is stage $n > N_f$ such that $\sigma_n \succ \sigma$, and a $\tau \succ \sigma_n$ such that $|v_\tau - f(q_{n-1})| \geq 1/k$. But then $I(\sigma) \not\subset \mathcal{P}$, a contradiction. Thus, $\hat{M}$ queries G only finitely often.

It remains to show that $\hat{M}$ infers f. Let $N > N_f$ be so large that the consequent of step 3 is not invoked after stage N, and fix $n > N$ (note that $\sigma_n = \sigma_N$). If $|a_n - f(q_n)| \geq 1/k$, then there are two cases to consider. If some $\tau \prec \sigma_N$ was used to compute a_n, then $a_n = M^G(\langle f(q_0), \ldots, f(q_{n-1})\rangle)$, contradicting that $n > N_f$. On the other hand, if some $\tau \succ \sigma_N$ was used, then at stage $n+1$ we have $v_\tau = a_n$, so that the consequent of step 3 is invoked at this stage, contradicting that $n > N$. □

Definition 6.1.6. Let M be any inference machine which outputs indices. The *amalgamation* procedure, AM, is defined from $M, \sigma, \langle x_0, \ldots, x_{n-1}\rangle$ as follows:

Case 1: $(M^\sigma(\langle x_0, \ldots, x_{n-1}\rangle)\downarrow)$. Output $M^\sigma(\langle x_0, \ldots, x_{n-1}\rangle)$.

Case 2: $(M^\sigma(\langle x_0, \ldots, x_{n-1}\rangle)\uparrow)$. For all $(M, \langle x_0, \ldots, x_{n-1}\rangle)$-minimal $\tau \succ \sigma$, compute $e_\tau = M^\tau(\langle x_0, \ldots, x_{n-1}\rangle)$. Output e defined by

$$\Phi_e(x) = \begin{cases} \Phi_{e_\tau}(x), & \text{for the first } \tau \text{ s.t. } \Phi_{e_\tau}(x)\downarrow, \\ \uparrow, & \text{if } \Phi_{e_\tau}(x)\uparrow \text{ for each } \tau. \end{cases}$$

End of Procedure.

THEOREM 6.1.7. If G is generic, then $BC_{1/k}[G*] = BC_{1/k}[G]$.

PROOF. Let $C \in BC_{1/k}[G]$ via M^G. We construct $\hat{M}$ to $BC_{1/k}[G*]$-infer C. Let $f \in \mathbb{Q}REC$, and initialize σ_0 to $\emptyset$. Then $\hat{M}$ works as follows:

Stage n:

Step 1: Output e_n, some uniformly chosen index of

$$AM(\sigma_n, \langle f(q_0), \ldots, f(q_{n-1})\rangle).$$

Step 2: For each $\tau \succ \sigma$, and each m with $|\sigma_n| < m < n$ such that τ is $(M, \langle f(q_0), \ldots, f(q_m)\rangle)$-minimal compute

$$e_\tau^m = M^\tau(\langle f(q_0), \ldots, f(q_m)\rangle),$$

and for each $x < n$, run $v_\tau^m(x) = \Phi_{e_\tau^m}(q_x)$ for n steps.

Step 3: If for some τ, m, and x we have $v_\tau^m(x)\downarrow$, with $|v_\tau^m(x) - f(q_x)| \geq 1/k$, let $\sigma_{n+1} = G|_n$. Otherwise, let $\sigma_{n+1} = \sigma_n$.

End of Construction.

As in 6.1.5, for $f \in \mathcal{C}$, $\hat{M}$ makes only finitely many queries to G. Since f is $BC_{1/k}$-inferred by M^G, there is an N_f so that for each $n > N_f$,

$$\|\Phi_{M^G(\langle f(q_0),\ldots,f(q_{n-1})\rangle)} - f\|_\infty < 1/k.$$

Let e_B^n denote $M^B(\langle f(q_0), \ldots, f(q_{n-1})\rangle)$. Let

$$\mathcal{P} = \{B \mid \forall n > N_f\ \forall m \text{ s.t. } (N_f < m < n)\ \forall x$$
$$\Phi_{e_B^m}^n(x)\downarrow \Rightarrow |\Phi_{e_B^m}^n(x) - f(x)| < 1/k\}.$$

Then $\mathcal{P}$ is a Π_1^0 class, and $G \in \mathcal{P}$, so there is a $\sigma \prec G$ with $I(\sigma) \subset \mathcal{P}$. Thus, the consequent of step 3 is invoked only finitely often, for otherwise, there is a stage $n > N_f$, $x < n$ and $\tau \succ \sigma_n$, with $\sigma_n \succ \sigma$, such that $|v_\tau(x) - f(q_x)| \geq 1/k$. But this contradicts that $I(\sigma) \subset \mathcal{P}$. Thus, $\hat{M}$ queries G only finitely often.

Finally, we show that $\hat{M}$ infers f. Let $\hat{\sigma} = \lim_n \sigma_n$, $N > \max\{N_f, |\hat{\sigma}|\}$, and fix $n > N$. If $\|\Phi_{e_n} - f\|_\infty \geq 1/k$, then there is an x with $|\Phi_{e_n}(q_x) - f(q_x)| \geq 1/k$. Thus, by definition of AM, there is a $\tau \succ \hat{\sigma}$ and an s such that $\Phi^s_{e^n_\tau}(q_x)\downarrow$, but $|\Phi^s_{e^n_\tau}(q_x) - f(q_x)| \geq 1/k$. Then at stage $\hat{n} = \max\{n, s+1\}$, the consequent of step 3 is invoked, contradicting that $\hat{n} > N$. □

DEFINITION 6.1.8. Let $M^{()}$ be a categorical oracle T.M., let $A \in 2^\omega$, and let $\mathcal{C} \subset \mathbb{Q}REC$. $\varepsilon : I_{\mathbb{Q}}^{<\omega} \longrightarrow \omega$ is an *NV_∞-modulus* for the pair $\langle M^A, \mathcal{C}\rangle$ if $\mathcal{C}$ is NV_∞-inferred by M^A, and for each $f \in \mathcal{C}$, $\exists N\ \forall k > N\ \forall n$,

$$n > \varepsilon(\langle f(q_0), \dots, f(q_k)\rangle) \Rightarrow |M^A(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f(q_n)| < 1/k.$$

$\varepsilon : I_{\mathbb{Q}}^{<\omega} \longrightarrow I_{\mathbb{Q}}$ is a *BC_∞-modulus* for the pair $\langle M^A, \mathcal{C}\rangle$ if $\mathcal{C}$ is BC_∞-inferred by M^A, and for each $f \in \mathcal{C}$, $\exists N\ \forall k > N\ \forall n$,

$$n > \varepsilon(\langle f(q_0), \dots, f(q_k)\rangle) \Rightarrow \|\Phi_{M^A(\langle f(q_0), \dots, f(q_{n-1})\rangle)} - f\|_\infty < 1/k.$$

Without loss of generality, we may assume that ε is increasing.

In the sequel, for a fixed f, we will often abuse notation, denoting $\varepsilon(\langle f(q_0), \dots, f(q_n)\rangle)$ by $\varepsilon(n)$.

THEOREM 6.1.9. Let G be generic, and let $\mathcal{C} \in NV_\infty[G]$ via M^G. If M has a recursive NV_∞-modulus, then $\mathcal{C} \in NV_\infty[G*]$.

PROOF. Suppose M^G NV_∞-infers $\mathcal{C}$ with recursive modulus ε. We construct $\hat{M}$ to $NV_\infty[G*]$-infer $\mathcal{C}$. Let $f \in \mathbb{Q}REC$, initialize σ_0 to $\emptyset$, and denote by k_n the largest $k < n$ such that $n > \varepsilon(k)$. Note that the sequence $\{k_n\}$ is recursive, uniformly in f. Then $\hat{M}$ works as follows:

Stage n:

Step 1: Output $a_n = M^{\sigma_n 0^*}(\langle f(q_0), \dots, f(q_{n-1})\rangle)$.

Step 2: For each $\tau \succ \sigma_n$ such that τ is $(M, \langle f(q_0), \dots, f(q_{n-2})\rangle)$-minimal compute

$$v_\tau = M^\tau(\langle f(q_0), \dots, f(q_{n-2})\rangle).$$

Step 3: If for some τ we have $|v_\tau - f(q_{n-1})| \geq 1/k_n$, let $\sigma_{n+1} = G|_n$, otherwise, let $\sigma_{n+1} = \sigma_n$.

End of Construction.

Suppose $f \in \mathcal{C}$. We claim that $\hat{M}$ makes only finitely many queries to G. Since f is NV_∞-inferred by M^G, there is an K_f such that for each $k > K_f$ and each n,

$$n > \varepsilon(k) \Rightarrow |M^G(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f(q_n)| < 1/k.$$

Let

$$\mathcal{P}_k = \{B \mid \forall n > \varepsilon(k)\ |M^B(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f(q_n)| < 1/k\}.$$

Then $\mathcal{P}_k$ is a Π^0_1 class. Since $\mathcal{P} = \bigcap_{k>K_f} \mathcal{P}_k$ is an effective intersection of Π^0_1 classes, and $G \in \mathcal{P}$, there is a $\sigma \prec G$ with $I(\sigma) \subset \mathcal{P}$.

Thus, the consequent of step 3 is invoked only finitely often, for otherwise, there is a $k > K_f$, a stage $n > \varepsilon(k)$, with $|\sigma_n| > k$ so that $\sigma \prec \sigma_n$, and a $\tau \succ \sigma_n$, such that $|v_\tau(x) - f(q_x)| \geq 1/k$. But then $I(\sigma) \not\subset \mathcal{P}_k$, a contradiction, since $k > K_f$. Thus, $\hat{M}$ queries G only finitely often.

It remains to show that $\hat{M}$ infers f. Let $k > K_f$, and $N > \varepsilon(k)$ be so large that the consequent of step 3 is not invoked after stage N, and fix $n > N$ (note that $\sigma_n = \sigma_N$). If $|a_n - f| \geq 1/k$, then there are two cases to consider. If some $\tau \prec \sigma_N$ was used to compute a_n, then $a_n = M^G(\langle f(q_0), \dots, f(q_{n-1})\rangle)$, contradicting that $n > \varepsilon(k)$ and $k > K_f$. On the other hand, if some $\tau \succ \sigma_N$ was used, then $a_n = M^\tau(\langle f(q_0), \dots, f(q_{n-1})\rangle)$, so at stage $n+1$, we have that $|v_\tau - f(q_{n-1})| \geq 1/k$, so that the consequent of step 3 is invoked, again a contradiction. □

THEOREM 6.1.10. Let G be generic, and let $\mathcal{C} \in BC_\infty[G]$ via M^G. If M^G infers $\mathcal{C}$ with recursive BC_∞-modulus, then $\mathcal{C} \in BC_\infty[G*]$.

PROOF. Let M^G have modulus ε. We construct $\hat{M}$ to $BC_\infty[G*]$-infer C. Let $f \in \mathbb{Q}REC$, initialize σ_0 to $\emptyset$, and denote by k_n the largest $k < n$ such that $n > \varepsilon(k)$. Then $\hat{M}$ works as follows:

Stage n:

Step 1: Output e_n, some uniformly chosen index of

$$AM(\sigma_n, \langle f(q_0), \dots, f(q_{n-1})\rangle).$$

Step 2: For each $\tau \succ \sigma$, and each m with $|\sigma_n| < m < n$ such that τ is $(M, \langle f(q_0), \dots, f(q_m)\rangle)$-minimal compute

$$e_\tau^m = M^\tau(\langle f(q_0), \dots, f(q_m)\rangle),$$

and for each $x < n$, run $v_\tau^m(x) = \Phi_{e_\tau^m}(q_x)$ for n steps.

Step 3: If for some τ, m, and x we have $v_\tau^m(x)\downarrow$, with $|v_\tau^m(x) - f(q_x)| \geq 1/k_m$, let $\sigma_{n+1} = G|_n$. Otherwise, let $\sigma_{n+1} = \sigma_n$.

End of Construction.

Suppose $f \in \mathcal{C}$. We claim that $\hat{M}$ makes only finitely many queries to G. Since f is BC_∞-inferred by M^G, there is an K_f such that for each $k > K_f$ and each n,

$$n > \varepsilon(k) \Rightarrow \|\Phi_{M^G(\langle f(q_0),\dots,f(q_{n-1})\rangle)} - f\|_\infty < 1/k.$$

Let e_B^n denote $M^B(\langle f(q_0), \dots, f(q_{n-1})\rangle)$. Let

$$\mathcal{P}_k = \{B \mid \forall n > \varepsilon(k)\ \forall m \text{ s.t. } (\varepsilon(k) < m < n)\ \forall x$$

$$\Phi_{e_B^m}^n(x)\downarrow \Rightarrow |\Phi_{e_B^m}^n(x) - f(x)| < 1/k\}.$$

Then $\mathcal{P}_k$ is a Π_1^0 class. Since $\mathcal{P} = \bigcap_{k>K_f} \mathcal{P}_k$ is an effective intersection of Π_1^0 classes, and $G \in \mathcal{P}$, there is a $\sigma \prec G$ with $I(\sigma) \subset \mathcal{P}$.

Thus, the consequent of step 3 is invoked only finitely often, for otherwise, there is a $k > K_f$, a stage $n > \varepsilon(k)$, with $|\sigma_n| > k$ so that the consequent of step 3 is invoked at this stage. So, there is an $x < n$, and m with $|\sigma_n| < m < n$, such that $\sigma \prec \sigma_n$, and a $\tau \succ \sigma_n$, such that $|v_\tau^m(x) - f(q_x)| \geq 1/k_m$. But then $I(\sigma) \not\subset \mathcal{P}_{k_m}$, a contradiction, since $k_m \geq k > K_f$. Thus, $\hat{M}$ queries G only finitely often.

It remains to show that $\hat{M}$ infers f. Let $k > K_f$, and $N > \varepsilon(k)$ be so large that the consequent of step 3 is not invoked after stage N, and fix $n > N$ (note that $\sigma_n = \sigma_N$). If $\|\Phi_{e_n} - f\|_\infty \geq 1/k$, then for some x, $|\Phi_{e_n}(q_x) - f(q_x)| \geq 1/k$. There are two cases to consider. If some $\tau \prec \sigma_N$ was used to compute e_n, then $e_n = M^G(\langle f(q_0), \ldots, f(q_{n-1})\rangle)$, contradicting that $n > \varepsilon(k)$ and $k > K_f$. On the other hand, if some $\tau \succ \sigma_N$ was used, then by definition of AM, $\Phi_{e_n}(q_x) = \Phi_{e_\tau}^m(q_x)$ for some m. Now, if $m \geq n$, then we have $|\Phi^m_{M^\tau(\langle f(q_0),\ldots,f(q_{n-1})\rangle)}(q_x) - f(q_x)| \geq 1/k$, so that $I(\sigma) \not\subset \mathcal{P}_k$, a contradiction, whence $m < n$. But in this case, since $n - 1 > \varepsilon(k), wehavethatk_{n-1} \geq k$. Thus, since $e_\tau = M^\tau(\langle f(q_0), \ldots, f(q_{n-1})\rangle)$, at stage n we have $|v_\tau^{n-1}(x) - f(q_x)| \geq 1/k_{n-1}$, so that the consequent of step 3 is invoked, which is again a contradiction, since $n > N$. □

6.2 Inference with Recursive Moduli

DEFINITION 6.2.1. We say $\mathcal{C} \in MNV_\infty[A]$ if there is a categorical O.T.M. $G^{()}$ such that $\mathcal{C} \in NV_\infty[A]$ via G^A, with recursive NV_∞-modulus ε. We define $MBC_\infty[A]$ similarly. We say $\mathcal{C} \in MNV_\infty$ if $\mathcal{C} \in MNV_\infty[A]$ for some recursive A, and define and MBC_∞ similarly.

We show in the sequel that the classes defined above lie strictly between the corresponding standard and approximate inference classes.

THEOREM 6.2.2. $NV \subsetneq MNV_\infty$.

PROOF. Note that there is a *single* recursive function which works as an NV_∞-modulus for *all* NV-inference machines M, and $\mathcal{C} \in NV$, namely

$$\varepsilon(\langle a_0, a_1, \dots, a_{n-1} \rangle) = n.$$

Thus, $NV \subset MNV_\infty$. Now, for $f \in \mathbb{Q}REC$, define $\hat{f}$ by $\hat{f}(x) = f(x)/(x+1)$. Let $\mathcal{C} = \{\hat{f} \mid f \in \mathbb{Q}REC\}$. Then $\mathcal{C} \in MNV_\infty$ via M which outputs 0 on any input. The modulus function is given by

$$\varepsilon(\langle a_0, a_1, \dots, a_{n-1} \rangle) = n.$$

However, it is easy to see that $\mathcal{C} \notin NV$. Suppose $\mathcal{C} \in NV$ via M. Then we can use M to build a machine $\hat{M}$ as follows:

$$\hat{M}(\langle f(q_0), \dots, f(q_{n-1}) \rangle) = (n+1) * M(\langle \hat{f}(q_0), \dots, \hat{f}(q_{n-1}) \rangle).$$

It is easy to see that $\hat{M}$ NV-infers $\mathbb{Q}REC$. ◻

THEOREM 6.2.3. $BC \subsetneq MBC_\infty$.

PROOF. As with NV-inference,

$$\varepsilon(\langle a_0, a_1, \dots, a_{n-1} \rangle) = n.$$

works as an BC_∞-modulus for *all* BC-inference machines M, and $\mathcal{C} \in BC$. Thus, $BC \subset MBC_\infty$. For $f \in \mathbb{Q}REC$, define $\hat{f}$ by linear interpolation of the following points: $\hat{f}(1/n) = (1/n) * f(q_n)$, with $\hat{f}(0) = 0$, and $\hat{f}(1) = 1$. Let $\mathcal{C} = \{\hat{f} \mid f \in \mathbb{Q}REC\}$. Then $\mathcal{C} \subset RUC$, so $\mathcal{C} \in BC_\infty$ by Theorem 4.2.4 (via a machine which at stage n outputs the linear interpolation of the n inputs). A recursive modulus function is given by

$$\varepsilon(\langle a_0, a_1, \dots, a_{n-1} \rangle) = k,$$

where k is greatest such that $1/k \in \{q_0, q_1, \dots, q_{n-1}\}$. However, it is easy to see that $\mathcal{C} \notin BC$. Suppose $\mathcal{C} \in BC$ via M. We use M to build a machine to $\hat{M}$ defined as

follows: on input $\langle f(q_0), \ldots, f(q_{n-1}) \rangle$), $\hat{M}$ outputs an index of the function $\phi_n(x)$ defined by

$$\phi_n(q_m) = m * \Phi_{M(\langle \hat{f}(q_0), \ldots, \hat{f}(q_{n-1}) \rangle)}(1/m).$$

Then $\hat{M}$ BC-infers $\mathbb{Q}REC$. □

THEOREM 6.2.4. $MNV_\infty \subsetneq NV_\infty$.

PROOF. We suppose that RUC is NV_∞-inferred by M, with recursive NV_∞ modulus ε, for a contradiction. The proof uses a variation of the diagonalization construction used previously in Theorem 3.3.1: use M, ε to construct f in RUC for which the modulus ε fails infinitely often. We construct f as follows (set $f(0) = f(1) = 0$, and let $\hat{q}_0 = 0$):

Stage n:

Step 1: If $q_n \leq \hat{q}_n$, then $f(q_n)$ is already defined, so do nothing.

Step 2: Otherwise, let $v_n = M(\langle f(q_0), \ldots, f(q_{n-1}) \rangle)$, and let k be greatest such that $n > \varepsilon(\langle f(q_0), \ldots, f(q_k) \rangle)$.

Step 3: If $v_n \geq 1/k$, then we violate $\varepsilon(\langle f(q_0), \ldots, f(q_k) \rangle)$ by extending f by zero to q_n (i.e. set $f(q) = 0$ for all $q \in [\hat{q}_n, q_n]$). On the other hand, if $v_n < 1/k$, we violate $\varepsilon(\langle f(q_0), \ldots, f(q_k) \rangle)$ by extending f with a "hat" of height $\frac{2}{k}$. Thus, we define $f(q)$ for $q \in [\hat{q}_n, q_n]$ by linear interpolation of the three points $(\hat{q}_n, 0)$, $(\frac{(q_n - \hat{q})}{2}, \frac{2}{k})$, and $(q_n, 0)$.

Step 4: Set $\hat{q}_n = \max\{q_0, \ldots, q_n\}$.

End of Construction.

Then $f \in RUC$, but by the construction, ε fails infinitely often on f. □

THEOREM 6.2.5. $MBC_\infty \subsetneq BC_\infty$.

PROOF. We suppose that RUC is BC_∞-inferred by M, with BC_∞ modulus ε, for a contradiction. As above, we use M, ε to construct f in RUC for which the modulus ε fails infinitely often. We construct f as follows (set $f(0) = f(1) = 0$, and let $\hat{q}_0 = 0$):

Stage n:

Step 1: If $q_n \leq \hat{q}_n$ then $f(q_n)$ is already defined, so do nothing.

Step 2: Otherwise, $q_n > \hat{q}_n$. Let $\hat{f}$ be defined by

$$\hat{f}(q) = \begin{cases} f(q), & \text{if } q \leq \hat{q}, \\ 0, & \text{otherwise,} \end{cases}$$

and let k be greatest such that $n > \varepsilon(\langle f(q_0), \dots, f(q_k)\rangle)$.

Step 3: Dovetail computations of $\{\Phi_{M(\hat{f}|_i)}(q_n)\}_{i \geq n}$ until we find a convergent computation $v_n = \Phi_{M(\hat{f}|_{i_n})}(q_n)$.

Step 4: Let $r = \max\{q_n, q_{i_n}\}$. If $v_n \geq 1/k$, then we violate $\varepsilon(\langle f(q_0), \dots, f(q_k)\rangle)$ by extending f by zero to r (i.e. set $f(q) = 0$ for all $q \in [\hat{q}_n, r]$). On the other hand, if $v_n < 1/k$, we violate $\varepsilon(\langle f(q_0), \dots, f(q_k)\rangle)$ by extending f with a "hat" of height $\frac{2}{k}$. Thus, we define $f(q)$ for $q \in [\hat{q}_n, r]$ by linear interpolation of the three points $(\hat{q}_n, 0)$, $(\frac{(r-\hat{q})}{2}, \frac{2}{k})$, and $(r, 0)$.

Step 5: Set $\hat{q}_n = \max\{q_0, \dots, r\}$.

End of Construction.

Then $f \in RUC$, but by the construction, ε fails infinitely often on f. $\square$

6.3 Inference from Sets of Oracles

We now turn to a different question concerning the use of oracles in inductive inference, one not addressed in Fortnow et al. [9]. Let $\mu()$ denote the usual measure on 2^ω. We could then define the following inference notions:

DEFINITION 6.3.1. A class $\mathcal{C}$ of recursive rational functions is *next-value approximable almost surely* ($\mathcal{C} \in NV^1_\infty$) if

$$\mu(\{A \mid \mathcal{C} \in NV_\infty[A]\}) = 1.$$

And more generally, we could define an apparent hierarchy of notions NV^a for $0 \leq a \leq 1$ by $\mathcal{C} \in NV^a_\infty$ if

$$\mu(\{A \mid \mathcal{C} \in NV_\infty[A]\}) \geq a.$$

We shall see in the sequel that this hierarchy collapses, i.e. for all $a > 0$, $NV^a_\infty = NV_\infty$. In what follows, we will show that the analogous hierarchies collapse for the other notions of standard and approximate inference. The technique we employ is one similar to that used by Sacks [23] to show that

$$\mu(\{B \mid A \leq_T B\}) > 0 \Rightarrow A \equiv_T \mathbf{0}.$$

DEFINITION 6.3.2. Let $A, B \in 2^\omega$, and let $\mathcal{I}$ be any inference notion. Then A is *$\mathcal{I}$-trivial* if $\mathcal{I}[A] = \mathcal{I}$. $\mathcal{I}[A] \subseteq \mathcal{I}[B]$ is denoted $A \leq_{\mathcal{I}} B$. The *$\mathcal{I}$-cone* above A is the set of oracles $\{B \mid A \leq_{\mathcal{I}} B\}$.

THEOREM 6.3.3. Let $\mathcal{C} \subset REC$. If $\mu(\{A \mid \mathcal{C} \in NV[A]\}) > 0$, then $\mathcal{C} \in NV$.

PROOF. Let $\mathcal{S} = \{A \mid \mathcal{C} \in NV[A]\}$, and suppose $\mu(\mathcal{S}) > 0$ (note that $\mathcal{S}$ is measurable). Let $\mathcal{S}_n = \{A \mid \mathcal{C} \in NV[A] \text{ via } \phi^B_n\}$. Since $\mathcal{S} = \bigcup_n \mathcal{S}_n$, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a. Now, there is an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$. Furthermore, since $\mathcal{G}$ is open, there is a basis element $\mathcal{B} = \bigcup^n_{k=1} I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$.

Now, suppose $f \in \mathcal{C}$. Then we may stratify $\mathcal{S}_e$ according to how soon f is inferred, as follows. For each $N \in \omega$, let

$$\mathcal{S}_{e,N} = \{B \in \mathcal{S}_e \mid \forall n > N \; \phi^B_e(\langle f(0), \dots, f(n-1)\rangle) = f(n)\}.$$

Note that $\{\mathcal{S}_{e,N}\}_{N=0}^{\infty}$ is nondecreasing, with $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}$. Thus, for any $f \in \mathcal{C}$, there is an N_f such that for all $N > N_f$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}) < a$. Hence, for $N > N_f$, $\mu(\mathcal{S}_{e,N} \cap \mathcal{B}) > 4a$, and $\mu(\mathcal{B} - \mathcal{S}_{e,N}) < 2a$. Figure 6.1 illustrates this situation.

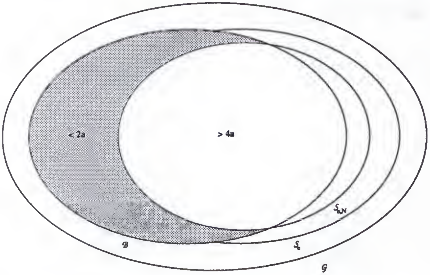

Figure 6.1: The relationship between $\mathcal{S}_{e,N}$ and $\mathcal{B}$ for $N > N_f$.

We now give the construction of M to NV-infer $\mathcal{C}$ (let $f \in \mathcal{C}$):

Stage n:

Step 1: Compute $v_B = \phi_e^B(\langle f(0), \ldots, f(n-1)\rangle)$ for all oracles $B \in \mathcal{B}$.

Step 2: Output the least value v for which $\mu(\{B \in \mathcal{B} \mid v_B = v\})$ is greatest.

End of Construction.

Since $\mathcal{B}$ is a finite union of intervals, only a finite number of computations are made in step 1 at any stage n. Not all of the ϕ_e^B necessarily NV-infer $\mathcal{C}$, but every computation in step 1 terminates, since $\phi_e^{()}$ is categorical. Thus, the algorithm is well-defined, and total for any input f. Now, if $f \in \mathcal{C}$, then for all stages $n > N_f$, we have $\mathcal{S}_{e,N_f} \subset \{B \in \mathcal{B} \mid v_B = f(n)\}$, so that $\mu(\{B \in \mathcal{B} \mid v_B = f(n)\}) > 4a$. Thus, since for each $n > N_f$, $\mu(\mathcal{B} - \mathcal{S}_{e,n}) < 2a$, M outputs $f(n)$ after stage N_f. $\square$

DEFINITION 6.3.4. The *majority amalgamation* procedure, *MAM*, is defined as follows. From parameters $I = \{e_1, e_2, \ldots, e_k\}$, and $W = \{w_1, w_2, \ldots, w_k\}$, on input x,

dovetail the computations $\phi_{e_1}(x), \ldots, \phi_{e_k}(x)$, until $e_{k_1}, e_{k_2}, \ldots, e_{k_n}$, are found such that each $\phi_{e_{k_i}}(x)$ converges to some common value v, and $\sum_{i=1}^{n} w_{k_i} > 1/2 \sum_{j=1}^{k} w_j$. Output v; if no such common value exists, then $MAM(I, W)(x) \uparrow$.

We give the analogous result for BC before the EX proof, since the BC case is easier.

THEOREM 6.3.5. Let $\mathcal{C} \subset REC$. If $\mu(\{A \mid \mathcal{C} \in BC[A]\}) > 0$, then $\mathcal{C} \in BC$.

PROOF. We define $\mathcal{S}$, $\mathcal{S}_e$, $\{\mathcal{S}_{e,N}\}_{N=0}^{\infty}$, $\mathcal{G}$ and $\mathcal{B}$ analogously to the sets of Theorem 6.3.3. Let $\mathcal{S} = \{A \mid \mathcal{C} \in BC[A]\}$, with $\mu(\mathcal{S}) > 0$. For each n, set $\mathcal{S}_n = \{A \mid \mathcal{C} \in BC[A] \text{ via } \phi_n^B\}$. As before, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a, an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$, and a basis element $\mathcal{B} = \bigcup_{k=1}^{n} I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$. If $f \in \mathcal{C}$, we stratify $\mathcal{S}_e$ according to how soon f is inferred, as follows. For each $N \in \omega$, let

$$\mathcal{S}_{e,N} = \{B \in \mathcal{S}_e \mid \forall n > N \ \phi_e^B(\langle f(0), \ldots, f(n-1)\rangle) \text{ is an index of } f\},$$

so that $\{\mathcal{S}_{e,N}\}_{N=0}^{\infty}$ is nondecreasing, with $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}$. Thus, for any $f \in \mathcal{C}$, there is an N_f such that for all $N > N_f$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}) < a$. Hence, for $N > N_f$, $\mu(\mathcal{S}_{e,N} \cap \mathcal{B}) > 4a$, and $\mu(\mathcal{B} - \mathcal{S}_{e,N}) < 2a$.

We construct M to BC-infer $\mathcal{C}$ as follows.

Stage n:

Step 1: Compute

$$I_n = \{e_{B,n} \mid e_{B,n} = \phi_e^B(\langle f(0), \ldots, f(n-1)\rangle), \text{ some } B \in \mathcal{B}\}.$$

Step 2: Enumerate the elements of I_n in increasing order $e_1, e_2, \ldots, e_k$. For each $i < k$, compute

$$w_i = \mu(\{B \in \mathcal{B} \mid \phi_e^B(\langle f(0), \ldots, f(n-1)\rangle) = e_i\}).$$

Denote the set of all w_i by $W(I_n)$.

Step 3: Output an index e_n (uniformly chosen) of $MAM(I_n, W(I_n))$.

End of Construction.

Analogously to Theorem 6.3.3, at every stage n, the computation in step 1 terminates, producing a finite set of indices I_n. If $f \in \mathcal{C}$, then for all stages $n > N_f$, we have

$$\mathcal{S}_{e,N_f} \subset \{B \in \mathcal{B} \mid \phi_e^B \text{ is an index of } f\},$$

so that for any x,

$$\mu(\{B \in \mathcal{B} \mid \phi_{e_{B,n}}(x) \downarrow = f(x)\}) > 4a.$$

Thus, for any $n > N_f$, $MAM(I_n, W(I_n)) \simeq f$. □

THEOREM 6.3.6. Let $\mathcal{C} \subset REC$. If $\mu(\{A \mid \mathcal{C} \in EX[A]\}) > 0$, then $\mathcal{C} \in EX$.

PROOF. We define $\mathcal{S}$, $\mathcal{S}_e$, $\{\mathcal{S}_{e,N}\}_{N=0}^{\infty}$, $\mathcal{G}$ and $\mathcal{B}$ analogously to the sets of Theorem 6.3.3. Let $\mathcal{S} = \{A \mid \mathcal{C} \in BC[A]\}$, with $\mu(\mathcal{S}) > 0$. For each n, set $\mathcal{S}_n = \{A \mid \mathcal{C} \in BC[A] \text{ via } \phi_n^B\}$. As before, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a, an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$, and a basis element $\mathcal{B} = \bigcup_{k=1}^{n} I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$. If $f \in \mathcal{C}$, we stratify $\mathcal{S}_e$ according to how soon f is inferred, as follows. For each $N \in \omega$, let

$$\mathcal{S}_{e,N} =$$
$$\{B \in \mathcal{S}_e \mid \exists e_B \, \forall n > N \; \phi_e^B(\langle f(0), \dots, f(n-1)\rangle) = e_B \text{ and } \phi_{e_B} \simeq f\},$$

so that $\{\mathcal{S}_{e,N}\}_{N=0}^{\infty}$ is nondecreasing, with $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}$. Thus, for any $f \in \mathcal{C}$, there is an N_f such that for all $N > N_f$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}) < a$. Hence, for $N > N_f$, $\mu(\mathcal{S}_{e,N} \cap \mathcal{B}) > 4a$, and so, $\mu(\mathcal{B} - \mathcal{S}_{e,N}) < 2a$.

Let $e_{B,n}$ denote $\phi_e^B(\langle f(0), \dots, f(n-1)\rangle)$. In the remaining theorems of this section, we will denote the set of indices produced by oracles $B \in \mathcal{A}$ for some $\mathcal{A} \subset 2^\omega$ at stage n as follows:

$$Ind(\mathcal{A}) = \{e_{B,n} \mid B \in \mathcal{A}\}.$$

Also, let $\mathcal{I}_0 = \emptyset$, and let $I_n = Ind(\mathcal{I}_n)$, and $F_n = Ind(\mathcal{F}_n)$. We construct M to EX-infer $\mathcal{C}$ as follows (let $n > 0$):

Stage n:

Step 1: Set $\mathcal{P}_n = \mathcal{I}_{n-1}$, and $\mathcal{F}_n = \{B \in \mathcal{B} \mid e_{B,n} = e_{B,n-1}\}$ (note that $\mathcal{F}_n$ is computable, since $\mathcal{B}$ is a finite union of intervals).

Step 2: If $\mu(\mathcal{P}_n \cap \mathcal{F}_n) \leq 4a$, let $\mathcal{I}_n = \mathcal{F}_n$, otherwise let $\mathcal{I}_n = \mathcal{P}_n$.

Step 3: Compute $W(I_n)$ as in Theorem 6.3.5.

Step 4: Output an index (uniformly chosen) of $MAM(I_n, W(I_n))$.

End of Construction.

Analogously to Theorem 6.3.3, at every stage n, the computation in step 1 terminates, producing a finite set of indices I_n. If $f \in \mathcal{C}$, then for all stages $n > N_f$, we have $\mathcal{S}_{e,N_f} \subset \mathcal{F}_n$, so that $\mu(F_n) > 4a$. Since $\mu(\mathcal{B} - \mathcal{S}_{e,n}) < 2a$ for $n > N_f$, for these n, the set of indices of f has relative weight greater than $1/2$ in F_n. Thus, for any $n > N_f$, $MAM(F_n, W(F_n)) \simeq f$.

We claim that for $n > N_f$, the consequent of step 2 is invoked at most once, and that $\lim_n M(\langle f(0), \ldots, f(n)\rangle)$ exists and is an index of f. If the consequent of step 2 *is* invoked after stage N_f, let $\hat{n}$ be the stage of the first such invocation. Then at this stage, we set $\mathcal{I}_{\hat{n}} = \mathcal{F}_{\hat{n}}$. Since $\hat{n} > N_f$, for all $n > \hat{n}$, $\mu(\mathcal{F}_n \cap \mathcal{F}_{\hat{n}}) > 4a$, so that the consequent of step 2 is not invoked at stage n, whence $\mathcal{I}_n = \mathcal{F}_{\hat{n}}$ for all $n > \hat{n}$. But then, as noted previously, $MAM(I_n, W(I_n)) \simeq f$, so M EX-infers f.

Suppose, on the other hand, that the consequent of step 2 is never invoked at any stage $n > N_f$. Then there is a (computable) set of oracles $\mathcal{I}$ with $\mu(\mathcal{I}) > 4a$, so that for all $n > N_f$, $\mathcal{I}_n = \mathcal{I}$. But then, for all $n > N_f$, M outputs a canonical index of $MAM(Ind(\mathcal{I}), W(Ind(\mathcal{I})))$. We claim that $MAM(Ind(\mathcal{I}), W(Ind(\mathcal{I})))$ computes f. Fix $n > N_f$. By hypothesis, $\mu(\mathcal{F}_n \cap \mathcal{I}) > 4a$. And by choice of n, $\mu((\mathcal{B} - \mathcal{S}_{e,n}) \cap \mathcal{I}) < 2a$.

Thus, the set of indices of f in $Ind(\mathcal{I})$ has relative measure greater than $1/2$, so that $MAM(Ind(\mathcal{I}), W(Ind(\mathcal{I})))$ computes f. $\square$

DEFINITION 6.3.7. Let $A = \{a_1, \dots, a_k\}$ and $W = \{w_1, \dots, w_k\}$ be finite sets of rationals (W is called a *weight set* for A, and w_i is called the *weight* of a_i). Let $M > 0$. We say that $B \subset A$ is an *$1/M$-majority set for* (A, W) if $\sum_{a_i \in B} w_i > 1/2 \sum_{i=1}^{k} w_i$, with $\operatorname{diam}(B) < 1/M$. We say that $1/M$ is the *majority number* of A if A has a $1/M$-majority set B, but no $1/N$-majority set for any $N > M$. If $B \subset A$ is a $1/M$-majority set for all $M > 0$, then necessarily B is a singleton, and is unique (we call B a 0-majority set). In this case, we say that A has majority number 0. We say that B is a majority set for (A, W) if B is a ε-majority set for some $\varepsilon > 0$.

Let

$$MSET(A, W) = \cup\{B \subset A \mid B \text{ is an } \varepsilon\text{-majority set}\},$$

and let $MAJ(A, W) =$ least element of $MSET(A, W)$, where ε is the majority number of A.

We verify some properties of majority sets which will be useful in the next few theorems.

LEMMA 6.3.8. Let $A = \{a_1, \dots, a_k\}$ and $W = \{w_1, \dots, w_k\}$ be finite sets of rationals. If B_1, B_2 are majority sets for (A, W), then $B_1 \cap B_2 \neq \emptyset$.

PROOF. Suppose $B_1 \cap B_2 = \emptyset$. Then,

$$\begin{aligned}
\sum_{a_i \in A} w_i &\geq \sum_{a_i \in B_1 \cup B_2} w_i \\
&= \sum_{a_i \in B_1} w_i + \sum_{a_i \in B_1} w_i \\
&< 1/2 \sum_{i=1}^{k} w_i + 1/2 \sum_{i=1}^{k} w_i \\
&= \sum_{a_i \in A} w_i,
\end{aligned}$$

a contradiction. $\square$

LEMMA 6.3.9. Let A and W be as above. If $1/M$ is the majority number of A, then $\operatorname{diam}(MSET(A, W)) < 2/M$.

PROOF. Let $b_1, b_2 \in MSET(A, W)$. Then there are $1/M$-majority sets B_1 and B_2 with $b_1 \in B_1$ and $b_2 \in B_2$. By 6.3.8, there is a $c \in B_1 \cap B_2$. Then

$$|b_1 - b_2| \leq |b_1 - c| + |c - b_2| < 1/M + 1/M = 2/M,$$

as desired. $\square$

THEOREM 6.3.10. Let $\mathcal{C} \subset \mathbb{Q}REC$. If $\mu(\{A \mid \mathcal{C} \in NV_\infty[A]\}) > 0$, then $\mathcal{C} \in NV_\infty$.

PROOF. Let $\mathcal{S} = \{A \mid \mathcal{C} \in NV_\infty[A]\}$, and suppose $\mu(\mathcal{S}) > 0$ (note that $\mathcal{S}$ is measurable). Let $\mathcal{S}_n = \{A \mid \mathcal{C} \in NV_\infty[A] \text{ via } \Phi_n^B\}$. Since $\mathcal{S} = \bigcup_n \mathcal{S}_n$, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a. Now, there is an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$. Furthermore, since $\mathcal{G}$ is open, there is a basis element $\mathcal{B} = \bigcup_{k=1}^n I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$.

Now, suppose $f \in \mathcal{C}$, and fix $M > 0$. Then we may stratify $\mathcal{S}_e$ according to how soon f is inferred to within $1/M$, as follows. For each $N \in \omega$, let

$$\mathcal{S}_{e,N}(M) = \{B \in \mathcal{S}_e \mid \forall n > N\ |\Phi_e^B(\langle f(q_0), \ldots, f(q_{n-1})\rangle) - f(q_n)| < 1/M\}.$$

Note that $\{\mathcal{S}_{e,N}(M)\}_{N=0}^\infty$ is nondecreasing, with $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}(M)$. Thus, for any $f \in \mathcal{C}$, $M > 0$, there is an $N_f(M)$ such that for all $N > N_f(M)$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}(M)) < a$. Hence, for $N > N_f(M)$, $\mu(\mathcal{S}_{e,N}(M) \cap \mathcal{B}) > 4a$, and $\mu(\mathcal{B} - \mathcal{S}_{e,N}(M)) < 2a$.

We now give the construction of G to NV_∞-infer $\mathcal{C}$. Let $f \in \mathbb{Q}REC$. Then G operates on f as follows.

Stage n:

Step 1: Compute $A_n = \{v \mid \exists B \in \mathcal{B}\ \Phi_e^B(\langle f(q_0), \ldots, f(q_{n-1})\rangle) = v\}$.

Step 2: Enumerate A_n as $v_1, \ldots, v_k$.

Step 3: Compute the weight set $W(A_n)$ for A_n:

$$w_i = \mu(B \in \mathcal{B} \mid \Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle) = v_i).$$

Step 4: Output $MAJ(A_n, W(A_n))$.

End of Construction.

Since $\mathcal{B}$ is a finite union of intervals, only a finite number of computations are made in step 1 at any stage n. Not all of the Φ_e^B necessarily NV_∞-infer $\mathcal{C}$, but every computation in step 1 terminates, since $\Phi_e^{()}$ is categorical. Thus, the algorithm is well-defined, and total for any input f. Now, if $f \in \mathcal{C}$ and $M > 0$, then for all stages $n > N_f(M)$, we have

$$\mathcal{S}_{e,N_f(M)}(M) \subset \{B \in \mathcal{B} \mid |\Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f(n)| < 1/M\},$$

so that

$$\mu(\{B \in \mathcal{B} \mid |\Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f(n)| < 1/M\}) > 4a.$$

Thus, since for each $n > N_f(M)$, $\mu(\mathcal{B} - \mathcal{S}_{e,n}(M)) < 2a$, so that the majority number of A_n is no greater than $2/M$. So, by the previous lemmas, G outputs values within $4/M$ of $f(n)$ after stage $N_f(M)$. □

DEFINITION 6.3.11. The *rational majority amalgamation* (*RMAM*) procedure is defined, for arbitrary input x, from parameters $\mathcal{I} = \{e_1, e_2, \dots, e_k\}$ and $W = \{w_1, w_2, \dots, w_k\}$ as follows.

Input x:

Step 1: Dovetail the computations of $\Phi_{e_1}(x), \dots, \Phi_{e_k}(x)$ until indices $e_{k_1}, \dots, e_{k_n}$ are found such that every $\Phi_{e_{k_i}}(x)\downarrow$, and

$$\sum_{i=1}^{n} w_{k_i} > 2/3 \sum_{j=1}^{k} w_j$$

(if no such indices are found, then $RMAM(\mathcal{I}, W)(x)\uparrow$).

Step 2: Enumerate the set $\{v \mid \exists i\ \Phi_{e_{k_i}}(x) = v\}$ as $\{v_1, \dots, v_l\}$.

Step 3: For each $i < l$, let $E_i = \{e_{k_j} \mid \Phi_{e_{k_j}}(x) = v_i\}$, and $r_i = \sum_{e_{k_j} \in E_i} w_{k_j}$.

Step 4: Output $MAJ(\{v_1, \dots, v_l\}, \{r_1, \dots, r_l\})$.

End of Procedure.

THEOREM 6.3.12. Let $\mathcal{C} \subset \mathbb{Q}REC$. If $\mu(\{A \mid \mathcal{C} \in BC_\infty[A]\}) > 0$, then $\mathcal{C} \in BC_\infty$.

PROOF. We define $\mathcal{S}$, $\mathcal{S}_e$, $\{\mathcal{S}_{e,N}(M)\}_{N=0}^{\infty}$, $\mathcal{G}$ and $\mathcal{B}$ analogously to the sets of Theorem 6.3.10. Let $\mathcal{S} = \{A \mid \mathcal{C} \in BC_\infty[A]\}$, with $\mu(\mathcal{S}) > 0$. For each n, set $\mathcal{S}_n = \{A \mid \mathcal{C} \in BC[A] \text{ via } \Phi_n^B\}$. As before, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a, an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$, and a basis element $\mathcal{B} = \bigcup_{k=1}^{n} I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$. If $f \in \mathcal{C}$, we stratify $\mathcal{S}_e$ according to how soon f is inferred, as follows. For each $N \in \omega$ and $M > 0$, let

$$\mathcal{S}_{e,N}(M) = \{B \in \mathcal{S}_e \mid \forall n > N\ \|\Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle) - f\|_\infty < 1/M\},$$

so that for fixed M, $\{\mathcal{S}_{e,N}(M)\}_{N=0}^{\infty}$ is nondecreasing (note that $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}(M)$). Thus, for any $f \in \mathcal{C}, M > 0$, there is an $N_f(M)$ such that for all $N > N_f(M)$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}(M)) < a$. Hence, for $N > N_f(\mathrm{M})$, $\mu(\mathcal{S}_{e,N}(M) \cap \mathcal{B}) > 4a$, and $\mu(\mathcal{B} - \mathcal{S}_{e,N}(M)) < 2a$. We construct G to BC_∞-infer $\mathcal{C}$ as follows.

Stage n:

Step 1: Compute

$$I_n = \{e_{B,n} \mid e_{B,n} = \Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle), \text{ some } B \in \mathcal{B}\}.$$

Step 2: Compute $W(I_n)$ as in Theorem 6.3.5.

Step 3: Output an index (uniformly chosen) of $RMAM(I_n, W(I_n))$.

End of Construction.

Analogously to Theorem 6.3.10, at every stage n, the computation in step 1 terminates, producing a finite set of indices I_n. If $f \in \mathcal{C}, M > 0$, then for all stages $n > N_f(M)$, we have

$$\mathcal{S}_{e,N_f(M)}(M) \subset \{B \in \mathcal{B} \mid \|\Phi_e^B - f\|_\infty < 1/M\},$$

so that for any x,

$$\mu(\{B \in \mathcal{B} \mid |\Phi_{e_{B,n}}(x) - f(x)| < 1/M\}) > 4a.$$

Thus, by Lemma 6.3.9, we have $\|RMAM(I_n, W(I_n)) - f\|_\infty < 3/M$ for any $n > N_f(M)$. □

DEFINITION 6.3.13. Let $f, g = \Phi_e \in \mathbb{Q}REC$. We say that g is an EX_∞-*variant* (and that e is an EX_∞-*index*) of f if

$$\lim_{n \to \infty} |g(q_n) - f(q_n)| = 0.$$

In this case we write $g \simeq_\infty f$.

THEOREM 6.3.14. Let $\mathcal{C} \subset \mathbb{Q}REC$. If $\mu(\{A \mid \mathcal{C} \in EX_\infty[A]\}) > 0$, then $\mathcal{C} \in EX_\infty$.

PROOF. We define $\mathcal{S}$, $\mathcal{S}_e$, $\{\mathcal{S}_{e,N}\}_{N=0}^\infty$, $\mathcal{G}$ and $\mathcal{B}$ analogously to the sets of Theorem 6.3.6. Let $\mathcal{S} = \{A \mid \mathcal{C} \in BC[A]\}$, $\mu(\mathcal{S}) > 0$. For each n, set $\mathcal{S}_n = \{A \mid \mathcal{C} \in EX_\infty[A]$ via $\Phi_n^B\}$. As before, there is an e with $\mu(\mathcal{S}_e) > 7a$ for some positive rational a, an open set $\mathcal{G}$ containing $\mathcal{S}_e$, with $\mu(\mathcal{G} - \mathcal{S}_e) < a$, and a basis element $\mathcal{B} = \bigcup_{k=1}^n I(\sigma_k)$ contained in $\mathcal{G}$ with $\mu(\mathcal{G} - \mathcal{B}) < a$. If $f \in \mathcal{C}$, we stratify $\mathcal{S}_e$ according to how soon f is inferred, as follows. For each $N \in \omega$, let

$$\mathcal{S}_{e,N} = \{B \in \mathcal{S}_e \mid \exists e_B\, \forall n > N\ \Phi_e^B((\langle f(q_0), \ldots, f(q_{n-1})\rangle)) = e_B \text{ and}$$
$$\Phi_{e_B} \simeq_\infty f\},$$

so that $\{\mathcal{S}_{e,N}\}_{N=0}^\infty$ is nondecreasing, with $\mathcal{S}_e = \bigcup_N \mathcal{S}_{e,N}$. Thus, for any $f \in \mathcal{C}$, there is an N_f such that for all $N > N_f$, $\mu(\mathcal{S}_e - \mathcal{S}_{e,N}) < a$. Hence, for $N > N_f$, $\mu(\mathcal{S}_{e,N} \cap \mathcal{B}) > 4a$, and so, $\mu(\mathcal{B} - \mathcal{S}_{e,N}) < 2a$.

Let $e_{B,n}$ denote $\Phi_e^B(\langle f(q_0), \dots, f(q_{n-1})\rangle)$, and let $\mathcal{I}_0 = \emptyset$. We construct G to EX_∞-infer $\mathcal{C}$ as follows (let $n > 0$):

Stage n:

Step 1: Set $\mathcal{P}_n = \mathcal{I}_{n-1}$, and $\mathcal{F}_n = \{B \in \mathcal{B} \mid e_{B,n} = e_{B,n-1}\}$.

Step 2: If $\mu(\mathcal{P}_n \cap \mathcal{F}_n) \leq 4a$, let $\mathcal{I}_n = \mathcal{F}_n$, otherwise let $\mathcal{I}_n = \mathcal{P}_n$.

Step 3: Compute $W(I_n)$ as above.

Step 4: Output an index (uniformly chosen) of $RMAM(I_n, W(I_n))$.

End of Construction.

Analogously to Theorem 6.3.6, at every stage n, the computation in step 1 terminates, producing a finite set of indices I_n. If $f \in \mathcal{C}$, then for all stages $n > N_f$, we have $\mathcal{S}_{e,N_f} \subset \mathcal{F}_n$, so that $\mu(\mathcal{F}_n) > 4a$. Fix $n > N_f$. Now, for each $M, K > 0$, define

$$\mathcal{T}_K(M) = \{B \in \mathcal{S}_{e,n} \mid \forall k > K\ |\Phi_e^B(q_k) - f(q_k)| < 1/M\}.$$

For fixed M, $\{\mathcal{T}_K(M)\}_{K=1}^{\infty}$ is an increasing sequence, with $\bigcup_{K=1}^{\infty} \mathcal{T}_K(M) = \mathcal{S}_{e,n}$. Thus, there is a $K_f(M)$ such that for all $k > K_f(M)$, $\mu(\mathcal{T}_k(M)) > 4a$. Thus, $\mu(\mathcal{B} - \mathcal{T}_k(M)) < 2a$ for $k > K_f$, so that for these k,

$$|RMAM(F_n, W(F_n))(q_k) - f(q_k)| < 2/M.$$

Thus, $RMAM(F_n, W(F_n)) \simeq_\infty f$.

We claim that for $n > N_f$, the consequent of step 2 is invoked at most once, and that $\lim_n G(\langle f(q_0), \dots, f(q_n)\rangle)$ exists and is an EX_∞-index of f. If the consequent of step 2 *is* invoked after stage N_f, let $\hat{n}$ be the stage of the first such invocation. Then at this stage, we set $\mathcal{I}_{\hat{n}} = \mathcal{F}_{\hat{n}}$. Since $\hat{n} > N_f$, for all $n > \hat{n}$, $\mu(\mathcal{F}_n \cap \mathcal{F}_{\hat{n}}) > 4a$, so that the consequent of step 2 is not invoked at stage n, whence $\mathcal{I}_n = \mathcal{F}_{\hat{n}}$ for all $n > \hat{n}$. But then, as noted previously, $RMAM(F_n, W(F_n)) \simeq_\infty f$, so G EX-infers f.

Suppose, on the other hand, that the consequent of step 2 is never invoked at any stage $n > N_f$. Then there is a set of oracles $\mathcal{I}$ with $\mu(\mathcal{I}) > 4a$ so that for all $n > N_f$, $\mathcal{I}_n = \mathcal{I}$, so that for all $n > N_f$, M outputs a canonical index of $RMAM(I, W(I))$, where I denotes $Ind(\mathcal{I})$. We claim that $RMAM(I, W(I)) \simeq_\infty f$, Fix $n > N_f, M > 0$, and let define $\{\mathcal{T}_K(M)\}$, and $K_f(M)$ as above. By hypothesis, $\mu(\mathcal{F}_n \cap \mathcal{I}) > 4a$. And by choice of n, for $k > K_f$(M), $\mu((\mathcal{B} - \mathcal{T}_k(M)) \cap \mathcal{I}) < 2a$. Thus, for these k, $|RMAM(I, W(I))(q_k) - f(q_k)| < 4/M$. Thus, $RMAM(I, W(I)) \simeq_\infty f$. □

The above results give us an analogue, for each inference class discussed, of the theorem of Sacks mentioned above:

Corollary 6.3.15. Let $\mathcal{I} \in \{NV, EX, BC, NV_\infty, EX_\infty, BC_\infty\}$. If A is not $\mathcal{I}$-trivial, then the $\mathcal{I}$-cone above A has measure zero.

Proof. Suppose $\mu(\{B \mid \mathcal{I}[A] \subseteq \mathcal{I}[B]\}) > 0$. Then, for any $C \in \mathcal{I}[A]$, $\mu(\{B \mid C \in \mathcal{I}[B]\}) > 0$, whence $C \in \mathcal{I}$. □

CHAPTER 7
SUMMARY AND CONCLUSIONS

7.1 Summary and Open Problems

We have utilized the standard metric on $\mathbb{Q}$ to extend the basic notions of inductive inference in a natural way, allowing us to infer a larger class of functions, and in particular, to infer classes of continuous functions. We have explored the relationships among these new notions of approximate inference, as well as between these notions and the basic notions NV, EX, and BC. Specifically, we gave precise inclusions between the new inference notions and those in the standard inference hierarchy. We also explored weaker notions of approximate inference, leading to inference hierarchies analogous to the EX^n and BC^n hierarchies. Oracle inductive inference was also considered, and we gave sufficient conditions under which approximate inference from a generic oracle G is equivalent to approximate inference with only finitely many queries to G. Whether these conditions are also necessary remains an open question. Finally, we employed approximation techniques from topology and analysis to obtain a new result regarding triviality in oracle inductive inference classes.

We have only begun to explore the area of approximate inductive inference. In the remaining sections, we offer some ideas for futher research in this field.

7.2 Stability

Recall that the standard inference hierarchy is linear, that is

$$NV = PEX \subsetneq EX \subsetneq BC,$$

but the analogous relation does not hold for the approximate inference classes. In particular, it is the class NV_∞ which "ruins" the analogy. We wish to explore ways

to redefine the notion of NV_∞ to remedy this situation. To show that $NV \subset PEX$, at stage N, one uses the outputs from an NV-machine M at stages $n > N$ as the inputs to the following stages to create a function to use as a guess for the input function at the stage N. We recall from chapter 3 the following definition (modified to accommodate functions on $I_{\mathbb{Q}}$), which formalizes this procedure.

DEFINITION 7.2.1. For any M, $\sigma = \langle a_0, \dots, a_{m-1}\rangle$, we define $S_{M,\sigma}$ on $I_{\mathbb{Q}}$ by recursion as follows

$$S_{M,\sigma}(q_n) = \begin{cases} M(\langle a_0, \dots, a_{n-1}\rangle), & \text{if } n < m, \\ M(\langle a_0, \dots, a_{m-1}, S_{M,\sigma}(q_m), \dots, \\ \quad S_{M,\sigma}(q_{n-1})\rangle), & \text{otherwise.} \end{cases}$$

We now introduce our first notion of stability for NV.

DEFINITION 7.2.2. We say that $\mathcal{C}$ is NV_∞-stable if there is an M which NV_∞-infers $\mathcal{C}$, and for all $f \in \mathcal{C}$ there is a stage L so that for all stages $l > L$,

$$\lim_{n\to\infty} |S_{M,\sigma_l}(q_n) - f(q_n)| = 0,$$

where $\sigma_l = \langle f(q_0), \dots, f(q_l)\rangle$. Denote by SNV_∞ the class of all such $\mathcal{C}$.

It is then easy to see the following.

PROPOSITION 7.2.3. $SNV_\infty = PBC_\infty$.

This does not quite get us the desired inclusion, however, since $PEX_\infty \subsetneq PBC_\infty$. We need an even stronger notion of stability to achieve this.

DEFINITION 7.2.4. We say that $\mathcal{C}$ is NV_∞-superstable if there is an M which SNV_∞-infers $\mathcal{C}$, and for all $f \in \mathcal{C}$ there is a stage L so that for all stages $k, l > L$,

$$S_{M,\sigma_k} = S_{M,\sigma_l}.$$

Denote by $SSNV_\infty$ the class of all such $\mathcal{C}$.

We then obtain the desired result:

PROPOSITION 7.2.5. $SSNV_\infty = PEX_\infty$.

7.3 Other Notions of Approximate Inference

Another scheme for defining notions of approximate inference is one in the style of Egorov's Theorem. We desire our inference method to get "close" to the input function, except on a set of size ε.

DEFINITION 7.3.1. We say that M ENV_ε-infers $\mathcal{C}$ if there is a computable set $E \subset [0,1]$ with $\mu(E) < \varepsilon$ such that for each $f \in \mathcal{C}$,

$$\lim_{k\to\infty} |M(\langle f(q_0),\ldots,f(q_{n_k-1})\rangle) - f(q_{n_k})| = 0,$$

where $\{q_{n_k}\}$ denotes the subsequence of $\{q_n\}$ given by the elements $q_n \notin E$.

DEFINITION 7.3.2. We say that M EEX_ε-infers $\mathcal{C}$ if there is a computable set $E \subset [0,1]$ with $\mu(E) < \varepsilon$ such that for each $f \in \mathcal{C}$, there is an index e (of a total function) and a stage N such that for all $n > N$, $M(\langle f(q_0),\ldots,f(q_n)\rangle) = e$, and

$$\lim_{k\to\infty} |\Phi_e(q_{n_k}) - f(q_{n_k})| = 0,$$

where $\{q_{n_k}\}$ denotes the subsequence of $\{q_n\}$ given by the elements $q_n \notin E$.

DEFINITION 7.3.3. We say that M EBC_ε-infers $\mathcal{C}$ if there is a computable set $E \subset [0,1]$ with $\mu(E) < \varepsilon$ such that for each $f \in \mathcal{C}$,

$$\lim_{k\to\infty} \|\Phi_{M(\langle f(q_0),\ldots,f(q_{n-1})\rangle)} - f\|_\infty^E = 0,$$

where $\|\cdot\|_\infty^E$ indicates that the infimum is taken over $q \in I_\mathbb{Q} - E$.

These definitions yield inference hierarchies distinct from the previous ones.

7.4 Inference of Non-recursive Functions

As noted in the introduction, the criteria for sucessful inference in the standard classes *NV*, *EX*, and *BC* limit us to inference of *recursive* functions. In contrast, the ideas of approximate inference allow us to extend the notion of inductive inference to include non-recursive functions. For example, the linear interpolation procedure works to NV_∞- or BC_∞-infer the class of *all* uniformly continuous functions on $I_\mathbb{Q}$

(of which RUC is a proper subset). Now, if we take our domain of inference to be the set of all $f : I_{\mathbb{Q}} \longrightarrow I_{\mathbb{Q}}$, many interesting questions arise. Clearly, not all of these functions are inferable by any of the methods given. For example, fix any non-computable irrational a in $[0,1]$, then if we define $\chi : I_{\mathbb{Q}} \longrightarrow I_{\mathbb{Q}}$ by

$$\chi(q) \;=\; \begin{cases} 0, & \text{if } q < a, \\ 1, & \text{if } q > a, \end{cases}$$

then the singleton $\{\chi\}$ is not NV_∞- or BC_∞-inferable: since χ is $0,1$-valued, if it is NV_∞ (resp. BC_∞) inferable, then it is NV (resp. BC) inferable.

7.5 Inference of Real-valued Functions

Slightly generalizing the input procedure for the approximate inference classes will allow us to further extend the domain of inference to include all real-valued functions (for an alternate formulation see Apsitis, Freivalds and Smith [2]). The actual machinery is only slightly changed. Suppose that f maps $[0,1]$ into $[0,1]$. We assume some fixed enumeration $\{q_n\}$ of the elements of $I_{\mathbb{Q}}$, in which each rational appears infinitely often, and make guesses based on finite sequences of *pairs* of rationals $\langle q, r\rangle$, where r represents a rational approximation $f(q)$ for the input function f. We may then use these "updates" to the approximation of $f(x)$ for each x to try to NV_∞- or BC_∞-infer f. In fact, the usual linear interpolation procedure, modified to use at each stage the latest approximations given, works to infer the class of *all* continuous functions mapping $[0,1]$ into $[0,1]$. Since the continuous functions on $[0,1]$ are determined by their values on $I_{\mathbb{Q}} = [0,1]\cap\mathbb{Q}$, we only need to approximate $f(q)$ for rational q. But q appears infinitely often in our enumeration $\{q_n\}$, say as the subsequence $\{q_{n_k}\}$, so the interpolation procedure M will produce approximations $a_k = M(\langle f(q_0), \ldots, f(q_{n_k-1})\rangle)$, whence $f(q) = \lim_k a_k$.

7.6 Inductive Feature Extraction

We may wish, for instance, to compute f' or $\int f$ from f. All of the inductive inference paradigms, standard as well as approximate, can be used as "feature

extraction" tools to compute in this manner. For example, the standard inference classes can be used to compute "formal" derivatives of the class of polynomials over $\mathbb{N}$. With the techniques of approximate inference we can do a bit more. For elements of $\mathbb{Q}REC$, we can compute approximations to (true) derivatives. For example, if $\mathcal{C} \subset RUC$ is a class of functions f for which f' is continuous, we can use linear interpolation, along with the mean value theorem, to construct a machine M which, upon input $f(q_0), f(q_1), \ldots$ outputs functions which approximate f' in the NV_∞ or BC_∞ sense. Note that it is not necessarily the case that f' is an element of $\mathbb{Q}REC$, or that its range is contained in $I_{\mathbb{Q}}$. Thus, this type of feature extraction provides a natural setting in which to extend the domain of functions under consideration to ones which are non-recursive, and real-valued.

REFERENCES

[1] D. Angluin and C. H. Smith, *Inductive inference: theory and methods*, Computing Surveys 15, 1983, pp.237–269.

[2] K. Apsitis, R. Freivalds and C. H. Smith, *On the inductive inference of real-valued functions*, Proceedings of the Eighth Annual Conference on Computational Learning Theory, Association for Computing Machinery, 1995, pp.170–177.

[3] Barzdin, *Prognostication of automata and functions*, Information Processing '71, North Holland, 1972, pp.81–84.

[4] L. Blum and M. Blum, *Toward a mathematical theory of inductive inference*, Information and Control 28, 1975, pp.125–155.

[5] J. Case, S. Jain and A. Sharma, *On learning limiting programs*, International Journal of Foundations of Computer Science 3, 1992, pp.93–115.

[6] J. Case and C. Smith, *Comparison of identification criteria for machine inductive inference*, Theoretical Computer Science 25, 1983, pp.193–220.

[7] D. Cenzer and J. Remmel, Π^0_1*-classes in analysis*, (to appear).

[8] J. Feldman, *Some decidability results on grammatical inference and complexity*, Information and Control 20, 1972, pp.244–262.

[9] L. Fortnow, W. Gasarch, S. Jain, E. Kinber, M. Kummer,S. Kurtz, M. Pleszkoch, T. Slaman, R. Solovay, F. Stephan, *Extremes in the degrees of inferability* (Manuscript).

[10] E. M. Gold, *Language identification in the limit*, Information and Control 10, 1967, pp.447–474.

[11] P. Halmos, Measure Theory, Van Nostrand, New York, 1950.

[12] L. Harrington, unpublished manuscript.

[13] C. Jockusch, *Degrees of generic sets*, Recursion Theory: Its Generalizations and Applications, Cambridge University Press, New York, 1980, pp.110–139.

[14] M. Kummer and F. Stephan, *On the structure of degrees of inferability*, Proceedings of the Sixth Annual Conference on Computational Learning Theory, Morgan Kaufmann, San Mateo, CA, 1993, pp.117–126.

[15] J. Munkres, Topology, Prentice-Hall, Englewood Cliffs, NJ, 1975.

[16] P. Odifreddi, Classical Recursion Theory, Vol. I, North-Holland, Amsterdam, 1989.

[17] P. Odifreddi, *Inductive Inference* (Manuscript).

[18] M. B. Pour-El and J. I. Richards, Recursively Enumerable Sets and Degrees, Springer-Verlag, Berlin, 1989.

[19] K. Popper, The Logic of Scientific Discovery, Hutchinson, London, 1950.

[20] H. Royden, Real Analysis, Macmillan, New York, 1968.

[21] H. Rogers, Theory of Recursive Functions and Effective Computability, McGraw Hill, New York, 1967.

[22] J. Royer, *Inductive inference of approximations*, Information and Control 70, 1986, pp.156–178.

[23] G. Sacks, Degrees of Unsolvability, Springer-Verlag, Berlin, 1963.

[24] R. I. Soare, Recursively Enumerable Sets and Degrees, Springer-Verlag, Berlin, 1987.

BIOGRAPHICAL SKETCH

William Moser was born in Boston, Massachusetts, on March 17, 1962. He received bachelor's degrees in mathematics and computer science from the University of Massachusetts at Amherst in 1985. After working for three years as a programmer/analyst for M.I.T. Lincoln Laboratory, he returned to the University of Massachusetts, receiving an M.S. in mathematics in 1991. In 1996, he received a Ph.D. in mathematics from the University of Florida, Gainesville, FL, under the direction of Dr. Douglas Cenzer.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Douglas A. Cenzer

Douglas A. Cenzer, Chairman
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Joseph Glover
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William J. Mitchell
Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Rick L. Smith
Associate Professor of Mathematics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Joseph N. Wilson
Assistant Professor of Computer and Information Sciences

This dissertation was submitted to the Graduate Faculty of the Department of Mathematics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1996

Dean, Graduate School